8-14-2019 10:00 AM

Structural and Functional Characterization of Deinococcal DNA Damage Response A (DdrA)

Filip Todorovic
The University of Western Ontario

Supervisor
Junop, Murray S.
The University of Western Ontario

Graduate Program in Biochemistry
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

ABSTRACT

Deinococci exhibit a remarkable resilience toward DNA damage through the actions of several unique proteins, including DdrA. Although DdrA is critical for damage resistance, little is known about its mechanism of action. Despite sharing sequence similarity with Rad52, DdrA has been reported to lack single-stranded DNA annealing activity. In order to better characterize DdrA, structural studies were undertaken with the primary objective of gaining insight into the mechanism by which DdrA functions. Significant progress was made toward elucidating the X-ray crystal structure; in particular, identifying suitable DdrA domain boundaries for successful expression, purification and crystallization. In addition, we demonstrate for the first time that DdrA mediates ssDNA annealing to levels comparable to Rad52 *in vitro*. Residues (K22 and K105) critical for ssDNA binding and annealing were identified and further used to demonstrate that DdrA mediates resistance to extreme levels of DNA damage through its ability to anneal ssDNA *in vivo.*

Keywords: DdrA, *Deinococcus radiodurans*, DNA repair, X-ray crystallography, single-stranded DNA annealing, radioresistance, mitomycin C, cancer, extended synthesis-dependent strand annealing, Rad52, homologous recombination

SUMMARY FOR LAY AUDIENCE

D. radiodurans is a bacterium that was discovered in 1956 after surviving on a can of meat that had been exposed to intense doses of radiation. Typically, radiation kills organisms by destroying DNA, the molecule that all organisms require to maintain biological function. The fact that *D. radiodurans* was able to survive meant that the bacterium was either able to protect its DNA from damage in an extraordinary fashion or able to repair its DNA following damage in an extraordinarily efficient manner. Once it was established that both *D. radiodurans* and *E. coli*, a radiation-sensitive bacterium, accumulate DNA damage to the same extent, the latter was deemed to be true.

It was later determined that the unique resistance to radiation stems, in part, from a unique protein, known as DdrA. The protein was found to be “turned on” in the cell following DNA damage, providing correlative evidence that it is involved in repair. Furthermore, when the protein was eliminated from the cell, the organism became more radiation-sensitive, providing the first causative evidence that the protein is involved in repair. Prior to the publication of this thesis, it had been shown that DdrA is able to interact with DNA as one would expect for a protein required to repair DNA. However, beyond this, no further details regarding the exact role of DdrA were known. This thesis demonstrates that DdrA is capable of annealing DNA, which is an important aspect of genomic reconstruction following DNA damage. Furthermore, we have identified the exact regions of the protein, which are involved in this process. Most significantly, we have determined that without

this contribution by DdrA, cells are sensitive to DNA damage, underscoring the importance of this phenomenon in the living organism. To figure out exactly how annealing takes place at an atomic level, we are now interested in figuring out what DdrA bound to protein looks like using a technique, known as X-ray crystallography. Significant progress to this end has been made in the lab and our current research efforts are aimed at bringing this task to completion.

ACKNOWLEDGMENTS

This work would not have been possible without the guidance and supervision of Dr. Murray Junop, whose enthusiasm and round-the-clock availability have inspired my passion for science and facilitated my growth as a researcher. His scientific approach, stemming from fostering fundamental principles and attention to details, has had an influence not only on my scientific *modus operandi*, but also on my approach to life in general and for that I will be forever grateful.

I would like to further thank my advisors, Drs. Patrick O'Donoghue and Eric Ball, whose open-door policies have helped to enrich this journey from its inception all the way to the writing of this thesis. Their expertise in differing areas of biochemistry has been very helpful in overcoming unforeseen challenges.

I would also like to acknowledge all members of the Junop lab, past and present, most notably, Kun Zhang, Robert Szabla and Ryan Grainger for their willingness to help complete work needed to finish the thesis as well as Dr. Chris Brown for his invaluable mentorship. My journey would likewise not have been complete without the tireless efforts of my undergraduate students, Braeden Medeiros, Parisa Fani-Molky and Pallak Gupta.

Last but not least, I would like to thank my family – my parents, Dragan and Nataša, and my sister, Lana, whose unconditional love and unwavering support along the way have made my maturation into a scientist possible.

Table of Contents

Figures

Chapter 1: Introduction

Chapter 2: Functional Characterization of DdrA

Tables

Chapter 3: Structural Characterization of DdrA

Appendix

Abbreviations Used

6-4-PP: pyrimidine-(6-4)-pyrimidone photoproduct

APS: Advanced Photon Source

AUC: analytical ultracentrifugation

BER: base excision repair

BPP: bipyrimidine photoproduct

CPD: cyclobutyl pyrimidine dimer

D_{10}: dose yielding 10% survival

***D. radiodurans* or *D. rad*:** *Deinococcus radiodurans*

DAPI: 4′,6-diamidino-2-phenylindole

DdrA/B: DNA damage response A/B

DNA: deoxyribonucleic acid

DSB: double-strand break

dsDNA/ssDNA: double-stranded/single-stranded DNA

DTT: dithiothreitol

EDTA: ethylenediaminetetraacetic acid

EMSA: electrophoretic mobility shift assay

ESDSA: extended synthesis-dependent strand annealing

FPLC: fast protein liquid chromatography

FRET: Förster resonance energy transfer

H_2O_2: hydrogen peroxide

HR: homologous recombination

ICL: interstrand crosslink

IMAC: immobilized metal affinity chromatography

IPTG: isopropyl β-D-1-thiogalactopyranoside

IR: ionizing radiation

LB: lysogeny broth

MMC: mitomycin C

MMR: mismatch repair

MWCO: molecular weight cut-off

NER: nucleotide excision repair

NHEJ: non-homologous end joining

O_2^-: superoxide

$^1O_2$: singlet oxygen

OD_{600}: optical density 600

•OH: hydroxyl radical

PC: protein carbonylation

Phyre2: Protein **H**omology/Analog**y R**ecognition **E**ngine Version **2**.0

Pol I/III: DNA polymerase I/III

PSIPRED: Position-**S**pecific **I**terative-Basic Local Alignment Search Tool Based Secondary Structure **PRED**iction

ROS: reactive oxygen species

SAD: single-wavelength anomalous dispersion

SEC: size-exclusion chromatography

SEC-MALS: size-exclusion chromatography with multi-angle light scattering

SeMet: selenomethionine

SSA: single-strand annealing

SSAP: single-strand annealing protein

SSB: single-stranded DNA-binding protein

TBE: tris/borate/EDTA

TEV: tobacco etch virus

TGY: tryptone, glucose and yeast

UV: ultraviolet

Chapter 1:

Introduction

1.1 DNA Damage

Deoxyribonucleic acid (DNA) is the blueprint of life. DNA encodes all the genetic information required for constructing an organism and maintaining biological function. This information must therefore be faithfully protected from both endogenous as well as exogenous sources of damage.

DNA is constantly under attack from various types of damage. Oxidative damage is the most common type of stress that DNA encounters. This form of damage typically occurs through exposure to reactive oxygen species (ROS), which are a byproduct of many cellular metabolic processes. Reactive oxygen species can also be formed in cells as a result of exposure to exogenous agents, such as ionizing radiation (IR). Examples of ROS include, but are not limited to, hydrogen peroxide (H_2O_2), superoxide (O_2^-), singlet oxygen ($^1O_2$) and the hydroxyl radical (•OH). Base modification is the main form of oxidative damage induced by ROS with over 80 different types of modified bases having been documented (Bjelland & Seeberg, 2003). In addition, ROS can directly react with the sugar-phosphate backbone of DNA, resulting in single-strand breaks. If such breaks occur in close proximity, a DNA double-strand break (DSB) can be generated (Kozmin *et al.*, 2009). Since a single, unrepaired DSB is lethal to a cell, bacteria have evolved elaborate mechanisms to mitigate these risks. For instance, peroxidase, superoxide dismutase and catalase are all examples of enzymes capable of detoxifying ROS, thereby preventing DNA damage from occurring (Tian *et al.*, 2004). Additionally, in the event that DSB's do occur, several protein-driven

pathways have evolved to repair DSB's. In lower organisms, such as bacteria, DNA double strand breaks are typically repaired using the homologous recombination (HR) pathway.

Middle wave ultraviolet (UV-B 290-320 nm) and, in particular, short wave UV (UV-C 200-290 nm) radiation is further capable of inducing DNA damage, chiefly through the introduction of covalent linkages between DNA bases. In this manner, cyclobutyl pyrimidine dimers (CPD's) and pyrimidine-(6-4)-pyrimidone photoproducts (6-4-PP's) may be formed. Both of these lesions could prove to be lethal in bacteria if left unrepaired in large numbers, or mutagenic if bypassed during replication or repaired incorrectly (Pfeifer, 1997). Cells have evolved an arsenal of enzymes able to repair these lesions, including photolyases and DNA glycosylases as well as a multi-protein repair pathway known as nucleotide excision repair (NER). In addition to forming covalently linked DNA bases, ultraviolet radiation may also produce ROS, which could then damage DNA through strand breakage or base modification, as previously outlined.

Furthermore, chemical agents, produced in both the intra- and the extra-cellular environments, may form covalent bonds with DNA bases, producing adducts, which disrupt Watson-Crick base pairing. Disruption of base pairing may result in incorrect incorporation of bases during replication (De Bont & van Larebeke, 2004). Chemical adducts may also form interstrand crosslinks (ICL's), potentially inhibiting replication and transcription. Acrolein and malondialdehyde, produced by lipid oxidation, are two examples of mutagenic agents responsible for

the formation of ICL adducts. ICL's are also formed by clastogenic agents, most notably mitomycin C (MMC) and cisplatin (Crooke & Bradner, 1976). In bacteria, adducts and cross-links are most commonly repaired through the base excision repair (BER), NER and HR pathways.

1.2 *Deinococcus*

Bacteria of the genus *Deinococcus* are notoriously resistant to all forms of DNA damage and as such, the organism is of particular interest to the study of DNA repair. *Deinococcus radiodurans* (*D. radiodurans* or *D. rad*) was the first strain of the genus to be isolated. The bacteria were discovered in 1956 after surviving on a can of meat that had been exposed to 4 kGy of gamma radiation (Anderson *et al*., 1956). This is how the strain got its name, *radiodurans*, from the Latin "*radius*", meaning "ray of light", and "*durare*", meaning "to endure", altogether meaning "radiation resistant". It took many more years for *D. radiodurans* to be classified with other phylogenetically related bacteria into the *Deinococcus* genus (Brooks & Murray, 1981). To date, the genus includes 47 strains, all of which are remarkably resistant to a wide range of DNA damaging stimuli, including ionizing radiation, UV-C radiation, mitomycin C and desiccation (Battista, 1997).

1.2.1 General Features

D. radiodurans is a red-pigmented, non-pathogenic, Gram-positive bacterium. The organism is mesophilic, meaning that growth occurs most

optimally under moderate temperature conditions. *D. radiodurans* are easily cultured at 30°C in tryptone, glucose and yeast (TGY) media with a doubling time of approximately 2 hours. The complete genome of the bacterium is 3.28 Mb, consisting of two chromosomes (2.6 Mb and 0.4 Mb) and two plasmids (177.5 kb and 45.7 kb) (White *et al*., 1999). All four of these genetic elements are rich in protein-coding regions. In fact, across these four elements, 80.9-93.5% of the sequence encodes for protein. One third of the genes in *D. radiodurans* lack identifiable matches, suggesting that the organism encodes a particularly high amount of unique proteins, which, in part, explain the unique resistance of the bacterium to DNA damaging stimuli. In addition to this collection of unique proteins, genome analysis has identified homologues of proteins involved in well-characterized DNA repair pathways, such as mismatch repair (MMR), NER, BER and HR (Makarova *et al*., 2001). Interestingly, *D. radiodurans* continually maintains at least 2 (and as many as 10) complete copies of its genome. Although there are typically 4 copies, the exact number depends on the phase of growth and access to nutrients (Hansen, 1978). Having several genome copies is thought to aid in repair of DSB's by HR, although it has been reported that genome number does not appear to influence the degree of DNA damage resistance (Harsojo *et al*., 1981).

1.2.2 Ionizing Radiation

Perhaps the most impressive aspect of the DNA damage resistance profile of *D. radiodurans* is the resilience of the bacterium to ionizing radiation. Two

kGy of ionizing radiation is the dose yielding 10% survival (D_{10}) in *Escherichia coli* (*E. coli*), whereas the D_{10} for *D. radiodurans* is nearly 12 kGy. At 5 kGy, *D. radiodurans* exhibits practically no loss of viability. Interestingly, *D. radiodurans* does not resist exposure to DNA damage by protecting its genome prophylactically. In fact, both *D. radiodurans* and *E. coli* accumulate DNA damage to the same extent (Slade & Radman, 2011). A dose of 6 kGy results in the formation of approximately 200 double-strand breaks, 3,000 single-strand breaks and tens of thousands of altered bases in both organisms (Burrell *et al*., 1971). Therefore, it would appear that the ability of *D. radiodurans* to resist exceptionally large amounts of DNA damage is entirely due to the ability to restore its genome rapidly (Zahradka *et al*., 2006) and faithfully (Repar *et al*., 2010). The underlying DNA repair mechanisms responsible for this remarkable resistance to DNA damage are poorly understood.

1.2.3 Ultraviolet-C Radiation

D. radiodurans exhibits extraordinary resistance toward the DNA damaging properties of ultraviolet-C radiation. It is at least twenty times more resistant to UV-C than *E. coli*. While only ~40 J/m^2 of radiation are sufficient to kill 90% of *E. coli*, it takes more than 900 J/m^2 to achieve the same effect in *D. radiodurans* (Arrange *et al*., 1993). Notably, photolyases, which are capable of directly reversing the damage caused by UV-C, are absent in *D. radiodurans*. Likewise, the bacteria also lack functional SOS response machinery, which routinely repair the damage in *E. coli* (Makarova *et al*., 2001). Instead, UV-C induced damage in *D. radiodurans* is

repaired by a combination of NER (involving independent endonucleases, such as uvrA and uvsE) and recombinational repair (involving proteins, such as RecA, RecF and RecO) (Minton, 1994). It has been shown that a sub-lethal dose of approximately 500 J/m^2 results in the formation of tens of thousands of bipyrimidine photoproducts (BPP's), which are subsequently excised by uvrA and uvsE (Moeller *et al*., 2010) and released into the medium (Boling & Setlow, 1966). Amazingly, ~9% of the total genomic content of cells is released into the medium following exposure to this dose of radiation, corresponding to approximately 50 bases of DNA per BPP (Varghese & Day, 1970).

Higher doses of UV-C have been shown to induce extensive genomic fragmentation (Bonura & Smith, 1975). At high doses of radiation, large numbers of BPP's are formed. Subsequent excision of BPP's leaves many gaps in the DNA, which have the tendency to stall replication forks and lead to formation of DSB's. Furthermore, UV-C stimulates the production of ROS (Blaškovičová *et al*., 2017), which, as previously outlined, have the potential to form DBS's as well. It has been shown that inactivation of proteins required for recombinational repair (eg., RecA, RecO and RecF) renders *D. radiodurans* as sensitive to UV damage as mutations in uvrA and uvsE (Tanaka *et al*., 2005; Xu *et al*., 2008; Chang *et al*., 2010). This dependence on recombinational repair demonstrates the high degree of genomic fragmentation that occurs following UV exposure.

1.2.4 Mitomycin C

D. radiodurans is remarkably resistant to mitomycin C, a common chemotherapeutic agent that has been utilized for the treatment of a wide variety of cancers (Bradner, 2001). MMC is a cross-linking agent that forms deoxy- guanine monoadducts, dG-dG intrastrand crosslinks and dG-dG interstrand crosslinks (Weng *et al*., 2010). Excision of these adducts leads to the inhibition of transcription and, similarly to the excision of BPP's, the formation of double-strand breaks (Kitayama *et al*., 1983). At an MMC concentration of 1 µg/mL, *D. radiodurans* cultures experience no loss in viability after 40 minutes of exposure, whereas the same dose decreases survival in *E. coli* by three orders of magnitude by that time. Notably, *D. radiodurans* is also immune to the mutagenic effects of MMC that are commonly observed when *E. coli* is treated with sub-lethal doses of the drug (Sweet & Moseley, 1976). *D. radiodurans* is thought to respond to MMC- induced damage in a similar fashion to UV-C induced damage as RecA (Gutman *et al*., 1994) and uvrA (Moseley & Evans, 1983) mutant strains have been shown to exhibit sensitivity to MMC in addition to UV-C. These findings are unsurprising given the similarities in DNA damage induced by the two stimuli.

1.2.5 Desiccation

D. radiodurans is very resistant to extreme dryness, also referred to as desiccation. While only 0.1% of *E. coli* survive following 2 days of desiccation at <5% relative humidity, *D. radiodurans* remain fully viable after two weeks under the same conditions (Mattimore & Battista, 1996).

Given the scarcity of naturally occurring sources of ionizing radiation and the prevalence of deserts throughout the history of the Earth, the radioresistance of *Deinococcus* is believed to be a byproduct of adaption to periods of intense dehydration. To demonstrate this, 41 strains of *D. radiodurans* identified as being "radiation-sensitive" were found to be equivalently sensitive to desiccation (Mattimore & Battista, 1996). Furthermore, the correlation between resistance to IR and desiccation holds true in unrelated bacteria outside of the *Deinococcus* genus (Shukla *et al*., 2007). At a cellular level, the forms of damage observed in cells subjected to desiccation are similar to those observed in cells exposed to IR. Transcriptome analyses of *D. radiodurans* recovering from exposure to IR or desiccation revealed a subset of genes that respond similarly to both stimuli. Some of these genes encode conserved hypothetical proteins of unknown function, whereas other genes encode well characterized proteins involved in DNA maintenance and ROS scavenging (Tanaka *et al*., 2004).

Desiccation is capable of inducing DNA damage in three ways. First, the decrease in water availability leads to an increase in ROS production, which then damages DNA in ways that have been previously outlined (Section 1.1). Second, reduced water availability leads to protein denaturation. In this state, the function of DNA repair proteins is compromised, allowing for accumulation of various types of damage (Slade & Radman, 2011). Third, a dehydrated cell may enter cytostasis, whereby cellular processes become stagnant, further allowing DNA damage to accumulate until permissive growth conditions are restored (Potts, 1994). Depending on the length of time spent in cytostasis, a cell may accumulate large

numbers of differing types of DNA damage that must be repaired rapidly once growth is reinitiated. *D. radiodurans* has evolved the mechanisms necessary to meet this daunting challenge.

1.3 Factors Contributing to Resistance

D. radiodurans requires both protein and DNA synthesis for damage recovery. Correlation between radiation dose and repair kinetics suggests the existence of regulated checkpoints for DNA degradation, export, synthesis and replication. None of these phenomena are unique to *Deinococcus* and as such, are insufficient for explaining elevated radioresistance. Instead, three additional factors have been suggested to underlie *Deinococcal* damage resistance: physical scaffolding (Section 1.3.1), ROS scavenging (Section 1.3.2) and DNA repair (Section 1.4).

1.3.1 Physical Scaffolding

Initially, it was hypothesized that a peculiar toroidal (doughnut-shaped) arrangement of the genomic DNA, observed in stationary phase cells, may be responsible for the radioresistance of *D. radiodurans* (Levin-Zaidman *et al.*, 2003). It was thought that this condensed arrangement of the genome might help maintain proximity of broken ends through mechanical scaffolding. In this way, breaks could be rapidly resealed ‘in-place’, reducing the risk of joining wrong pairs of broken DNA segments. However, the absence of any genetic evidence for NHEJ being required for extreme DNA damage resistance in *Deinococcus* (Daly & Minton,

1996) as well as the dependence on recombinational repair (Daly *et al*., 1994), the increased radioresistance of cells cultivated in media inhibiting toroid formation (Daly *et al.*, 2004), and the fact that not all *Deinococci* store their genomic DNA in a toroidal conformation (Zimmerman & Battista, 2005) led to the demise of this hypothesis. Nevertheless, other methods of physical scaffolding, such as the formation of DNA-membrane complexes (Burrell *et al*., 1971) and the pre-alignment of homologous chromosomes (Minton & Daly, 1995) are still considered important factors contributing to DNA damage resistance in *Deinococci spp*.

1.3.2 Protection of the Proteome via ROS Scavenging

In response to the damaging effects of ROS, *Deinococcus spp.* have evolved a comprehensive array of 'protective' enzymes and free radical scavengers. In addition to export from the cell (as discussed in Section 1.2.3), damaged nucleotides are targeted for degradation by Nudix family hydrolases (Xu *et al*., 2001) and nucleotidases (Kota *et al*., 2010). *D. radiodurans* contains 23 different Nudix hydrolases, twice the number found in *E. coli*. Five of these hydrolases are upregulated following irradiation (Liu *et al*., 2003). In addition, *Deinococci spp*. maintain an expanded set of subtilisin-like proteases that serve to remove proteins that become modified, inactive or otherwise damaged during exposure to DNA damaging stimuli. The frequent removal of damaged proteins underscores the requirement for *de novo* protein synthesis prior to initiating DNA repair (Joshi *et al*., 2004).

While *D. radiodurans* accumulates DNA damage to the same degree as non-radiation resistant bacteria, the same is not true for the proteome of the bacterium, which is considerably better protected from oxidative damage (Daly *et al*., 2007; Krisko & Radman, 2010). Protein carbonylation (PC) is a common biomarker of oxidative stress that occurs when ROS react with amino acid side chains to generate reactive aldehydes and ketones (Dalle-Donne *et al.*, 2006). Unlike DNA damage, which accumulates to the same degree in *D. radiodurans* and *E. coli*, the rate of PC detected in *D. radiodurans* is 20 to 30 times lower than that in *E. coli* at equivalent doses of radiation (Krisko & Radman, 2010). A similar correlation between increased PC and decreased viability is, notably, observed in both organisms. The differential effects of oxidative damage to DNA and protein was inconsistent with the long-standing belief that ROS are capable of damaging macromolecules in an indiscriminate fashion. As seen in Figure 1.1, radiolysis of water leads to formation of three reactive oxygen species: •OH**,** O_2^- and H_2O_2. While all three have potential to inflict DNA damage, each species has differing downstream effects on protein and DNA. For instance, O_2^- is an inefficient oxidizing agent due to its negative charge and as such, does not act directly on DNA or amino acids (Imlay, 2003). Instead, O_2^- mainly targets iron-sulfur clusters without oxidizing the coordinating residues (Flint *et al*., 1993; Imlay, 2003). This form of oxidization could result in the termination of metabolic activity, since many of the metallo-proteins containing iron-sulfur clusters play key roles in metabolism and cellular respiration (Imlay *et al*., 2013). Likewise, H_2O_2 does not damage DNA, but will readily oxidize sulfur containing residues and residues coordinating either iron

or iron containing ligands (Imlay, 2003). •OH may be generated from the radiolysis of water or via the Fenton reaction and has the capacity to oxidize both protein and DNA. Unlike O_2^- and H_2O_2, which persist for long periods of time unless scavenged, •OH is short-lived and as such, can only react with molecules in the immediate vicinity (Imlay, 2008).

D. radiodurans possesses a wide arsenal of catalases, peroxidases and superoxide dismutases (see Section 1.1) for ROS neutralization (Makarova *et al*., 2001). Even though the relative activity of these enzymes is elevated in comparison to *E. coli*, *D. radiodurans* mutants deficient in catalase and superoxide dismutase activity were found to be only marginally more sensitive to the effects of ionizing radiation (Markillie *et al*., 1999). Furthermore, analysis of seven strains of *Deinococcus* found no correlation between elevated enzymatic scavenging of ROS and radioresistance (Shashidhar *et al*., 2010). These two findings suggest that the enzymes in question are primarily involved in the neutralization of ROS arising from ordinary cell activity and are of reduced importance in responding to ROS generated by external stimuli.

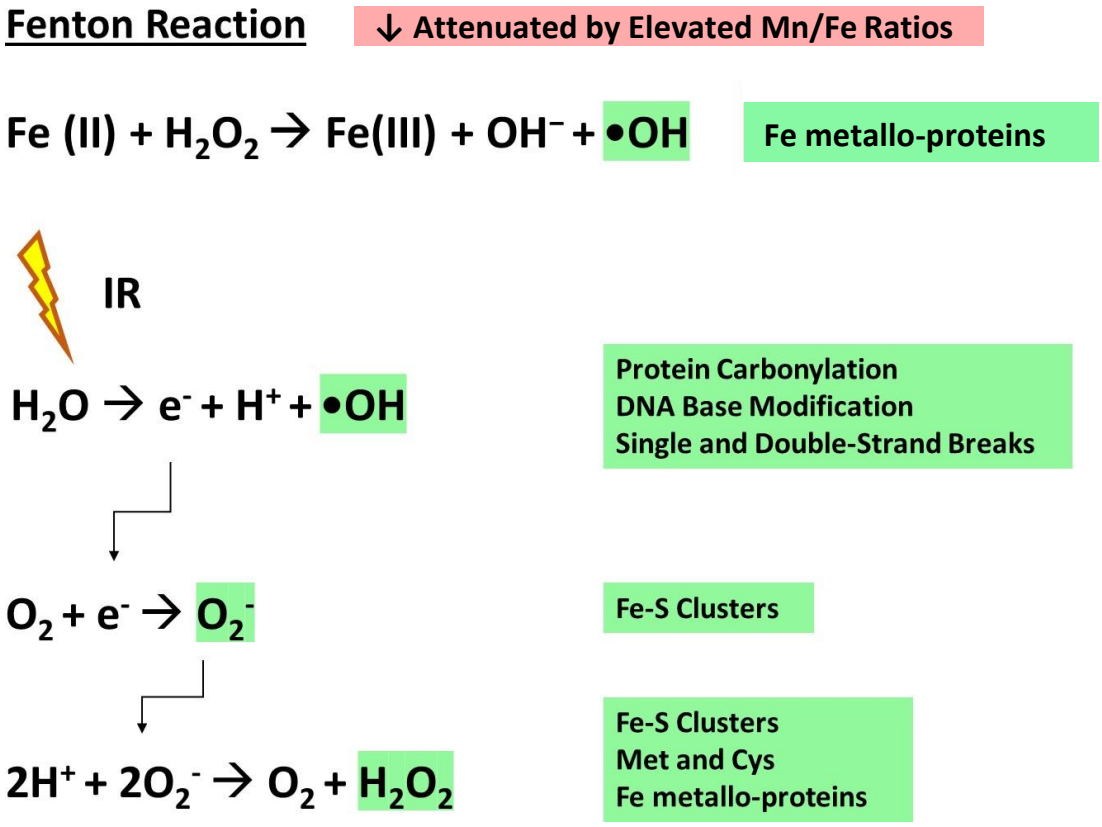


Figure 1.1: Generation of ROS. •OH may be formed either through the radiolysis of water or via the Fenton reaction, which is subject to attenuation by elevated manganese:iron ratios. •OH formed by Fenton chemistry primarily targets Fe metallo-proteins, whereas •OH formed via the radiolysis of water is responsible for PC, DNA base modification as well as single and double-strand break formation. The electron, which is generated as a byproduct of the radiolysis of water, is then free to react with molecular oxygen to form O_2^-, which chiefly targets iron-sulfur clusters. These superoxide radicals are then free to react with protons to form H_2O_2, which in addition to targeting iron-sulfur clusters, also oxidizes methionine, cysteine and residues coordinating either iron or iron containing ligands.

The main contributor to anti-oxidant activity is believed to be the heightened concentration of intracellular manganese (Mn) and manganese:iron ratio (Mn/Fe) found in all *Deinococci* as well other radioresistant organisms (Daly *et al*., 2007; Daly, 2009; Daly *et al*., 2010). *D. radiodurans* grown in TGY medium contains 0.36 nanomoles of manganese per milligram of protein and a Mn/Fe ratio of 0.24. In contrast, radiosensitive organisms such as *E. coli* and *Shewanella oneidensis* (*S. oneidensis*) have manganese concentrations of only 0.0197 nmol/mg and 0.0023 nmol/mg, respectively. Most significantly, when grown in media lacking Mn, the D_{10} of *D. radiodurans* drops from approximately 16 kGy of ionizing radiation to less than 2.5 kGy (Daly *et al*., 2004). An elevated Mn/Fe ratio is thought to be anti- oxidative in two ways. First, manganese complexes formed with orthophosphate and peptides act as efficient scavengers of H_2O_2 and O_2^- (Daly *et al*., 2010), the two ROS species most responsible for protein oxidation. Second, as illustrated in Figure 1.1, an elevated Mn/Fe ratio attenuates the Fenton reaction, thereby decreasing the amount of •OH that can be produced in this manner. This attenuation is protective of the proteome as •OH arising from Fenton chemistry is short-lived and since it is only produced in proximity of iron, it can therefore only target Fe metallo-proteins. In contrast, •OH formed by the radiolysis of water is indiscriminate.

The selective neutralization of reactive oxygen species that primarily target proteins explains the disproportionate protection of protein compared to DNA that is observed in *Deinococcus* following irradiation. It is possible that maintaining

the proteome in optimal working condition allows *Deinococcus* to more efficiently orchestrate the many protein-driven repair pathways necessary to respond to excessive DNA damage. Nevertheless, the lack of direct protection of DNA from ROS still necessitates an efficient mechanism for DNA repair. Since the work presented in this thesis relates to a protein (DdrA) involved in double-strand break repair, the following section will outline what is understood about these mechanisms in *Deinococcus*.

1.4 Recombinational Repair in *D. Radiodurans*

Massive fragmentation of the *Deinococcal* genome in response to DNA damage has been observed using pulsed-field gel electrophoresis (Grimsley *et al*., 1991). When bacteria are exposed to 7 kGy of ionizing radiation, a dose at which 90% of cells survive, the average size of DNA fragments is 20-30 kb, corresponding to approximately 100-200 double strand breaks per copy of the genome (Zahradka *et al*., 2006). Depending on the phase of growth and exact composition of the medium, *Deinococcus* may have up to 10 copies of the genome present. Since strand breakage occurs stochastically, the probability of the same locus being damaged in every single copy is thus negligible. Therefore, the cell always has a template from which to repair (Harsojo *et al*., 1981).

1.4.1 The RecBCD Pathway

The RecBCD complex is responsible for the initiation of homologous recombination in most bacteria and the majority of our understanding of this

pathway has been derived from studies of *E. coli* (Dillingham & Kowalczykowski, 2008). As illustrated in Figure 1.2, a double-strand break results in a free DNA end that is subsequently bound by the RecBCD complex. The RecBCD complex then unwinds the double-stranded DNA (dsDNA) and digests the resulting exposed single-stranded DNA (ssDNA) ends in a 3' to 5' fashion (Muskavitch & Linn, 1982). Degradation continues until the complex reaches a chi (χ) sequence (5'-GCTGGTGG-3'). At this point, the strand preference reverses (Anderson & Kowalczykowski, 1997). Exonuclease activity in the 5' to 3' direction becomes favoured and degradation of the 3' end is terminated. The 3' single-stranded DNA product is then coated by single-stranded DNA-binding protein (SSB) to prevent formation of secondary structures and further degradation (Muskavitch & Linn, 1982; Mackay & Linn, 1976). RecA is then recruited to this ssDNA/dsDNA junction by the RecBCD complex and polymerizes in 5' to 3' fashion, displacing SSB in the process. In this state, the resulting helical filament of RecA is primed for strand invasion of homologous duplex DNA (Tsang *et al*., 1985).

Unlike *E. coli*, *D. radiodurans* has no homologues of RecB and RecC. While it possesses a homologue of RecD, deletions of this gene do not result in radiosensitivity (Zhou et al., 2007). Furthermore, when RecBC from *E. coli* are expressed in *D. radiodurans*, radioresistance is not improved (Khairnar *et al*., 2008), suggesting that the initiation of homologous recombination in *D. radiodurans* occurs via a different mechanism.

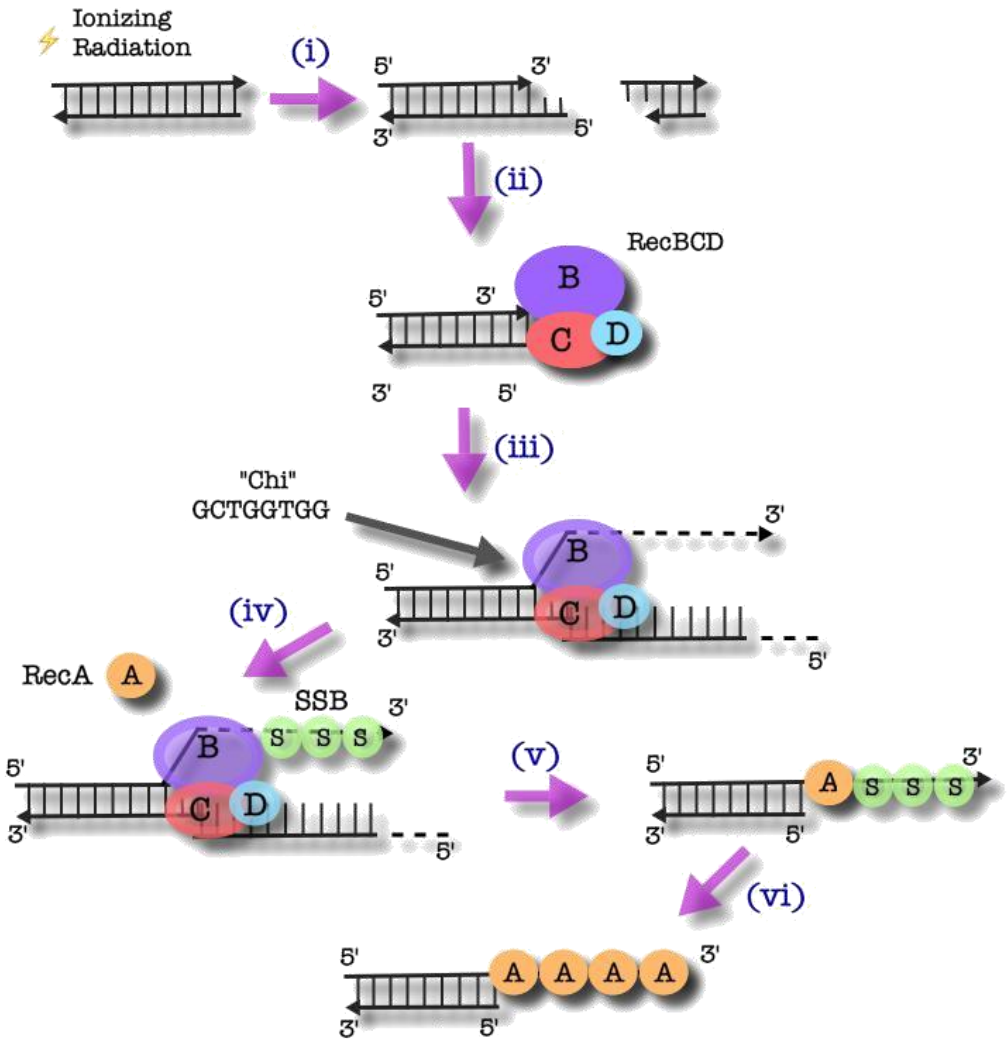


Figure 1.2: The Initiation of HR by the RecBCD Complex in *E. coli*

i) A double-strand break results from, in this instance, ionizing radiation.

ii) RecBCD complex binds the resulting DNA end.

iii) RecBCD complex unwinds the dsDNA and begins to digest the ssDNA, in a 3' to 5' fashion.

iv) When the complex reaches the χ region, degradation becomes favoured in the 5' to 3' direction, thus ending the breakdown of the 3' end. SSB coats the 3' strand to prevent secondary structure formation as well as further degradation.

v) RecA is recruited to the ss/dsDNA junction by the RecBCD complex.

vi) Polymerization of RecA in the 5' to 3' direction displaces SSB and forms a helical filament on the 3' ssDNA.

1.4.2 The RecFOR Pathway

In *E. coli*, RecBC-knockouts are capable of initiating the loading of RecA using the alternative RecFOR pathway (Lloyd & Buckman, 1985). *D. radiodurans* has homologues of the key proteins involved in the RecFOR alternative pathway (White *et al*., 1999). Furthermore, when sbcB, an inhibitor of the RecFOR pathway from *E. coli*, is expressed in *D. radiodurans,* a decrease in radioresistance is observed (Misra *et al*., 2006). Additionally, in *D. radiodurans,* RecF, RecO, RecR and RecA knockouts all exhibit radiosensitivity (Bentchikou *et al*., 2010), suggesting that the RecFOR pathway is the principal method for initiation of homologous recombination in *D. radiodurans*.

As illustrated in Figure 1.3, in the *Deinococcal* RecFOR pathway, uvrD recognizes broken DNA ends and unwinds the dsDNA. RecJ then degrades the ssDNA in a 5' to 3' fashion (Bentchikou *et al*., 2010). No sequence comparable to the χ region has been identified and as such, the mechanism by which the activity of RecJ is terminated remains uncertain. The other strand is then coated with either SSB or DNA damage response B (DdrB), a protein unique to *Deinococcus*, with no known homologues (Norais *et al*., 2009). RecFOR then binds the ss/dsDNA junction (Timmins *et al*., 2007). From here, RecA ultimately promotes strand exchange by homologous pairing of ssDNA and dsDNA by either coating a homologous duplex or the aforementioned 3' end that the SSB was coating.

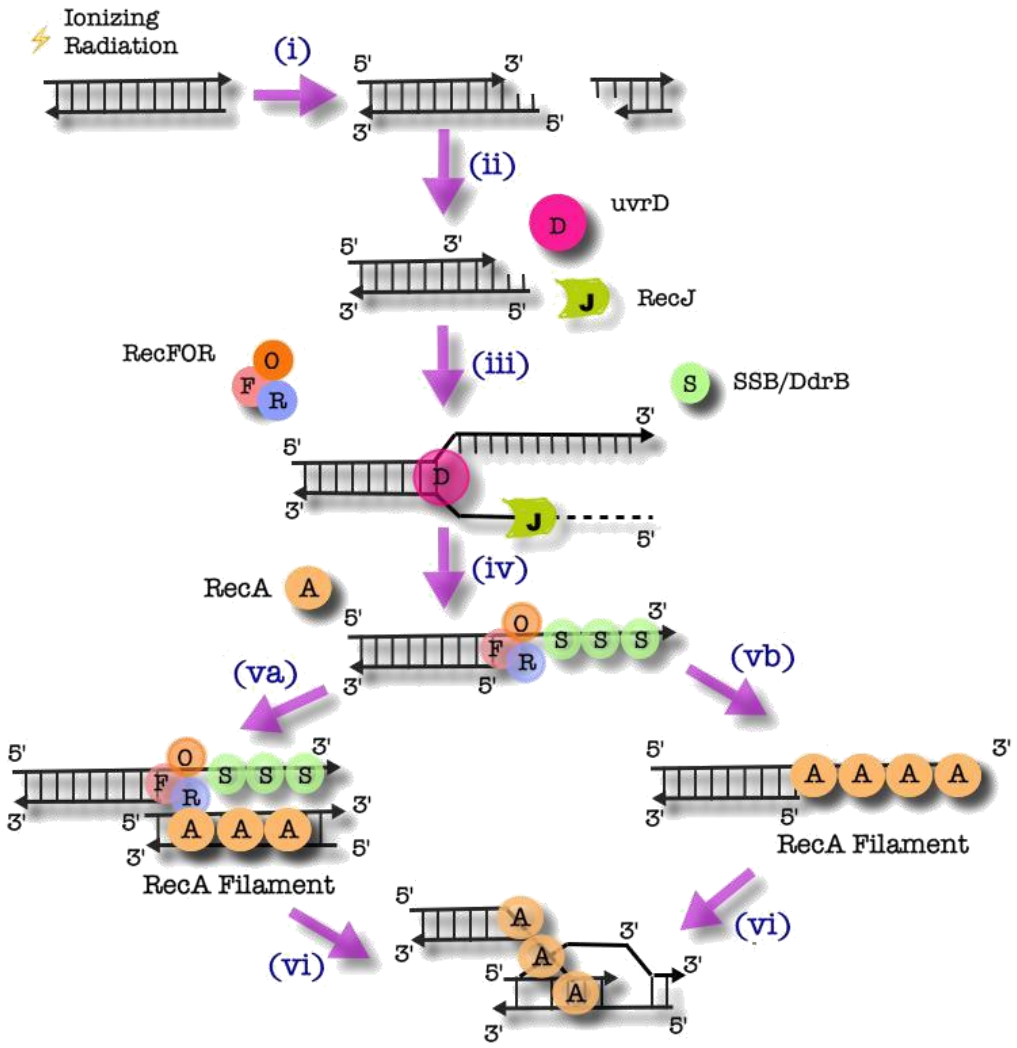


Figure 1.3: The Initiation of HR by the RecFOR Complex in *D. radiodurans*

i) A double-strand break is generated, in this example, by IR.

ii) For *D. radiodurans,* the proteins subsequently involved are uvrD and RecJ, whereas in *E. coli*, RecQ replaces uvrD.

iii) UvrD recognizes the DNA end resulting from a double-strand break and unwinds the dsDNA so that RecJ may degrade the ssDNA in a 5' to 3' fashion.

iv) SSB or DdrB then coat the opposing strand to protect from secondary structure formation and further degradation.

vi) RecA ultimately promotes strand exchange by homologous pairing of ssDNA and dsDNA by coating either va) a homologous duplex or vb) the 3' end that the SSB was previously coating.

1.4.3. Extended Synthesis-Dependent Strand Annealing

Since complete and efficient recovery from extensive genomic fragmentation requires significant DNA synthesis, conventional homologous recombination cannot suffice as the primary mechanism of DNA repair in *Deinococcus* (Zahradka *et al*., 2006). Instead, it is thought that a variation of recombinational repair, termed “extended synthesis-dependent strand annealing” (ESDSA) is used for repair of DSB’s.

According to this model (Figure 1.4), RecA-mediated strand invasion of homologous duplex DNA initially occurs to form a D-loop. Extension of the invading 3’ strand is carried out by DNA polymerase III (pol III). DNA polymerase I (pol I) is capable of facilitating, but not initiating, strand extension (Slade *et al*., 2009). The extended invading strand then disassociates from the template strand and either invades another homologous duplex to initiate a novel round of extension or anneals with a complementary extension formed via the same mechanism using a different template (Zahradka *et al*., 2006). Given the rapid speed by which long (up to 20 kb) ssDNA overhangs are converted to duplex DNA, a model has been proposed whereby extension occurs simultaneously along a single fragment that further serves as a bridge or a scaffold (Figure 1.5) (Slade *et al*., 2009). Finally, RecA-mediated recombination facilitates recircularization of the newly formed duplex DNA fragments.

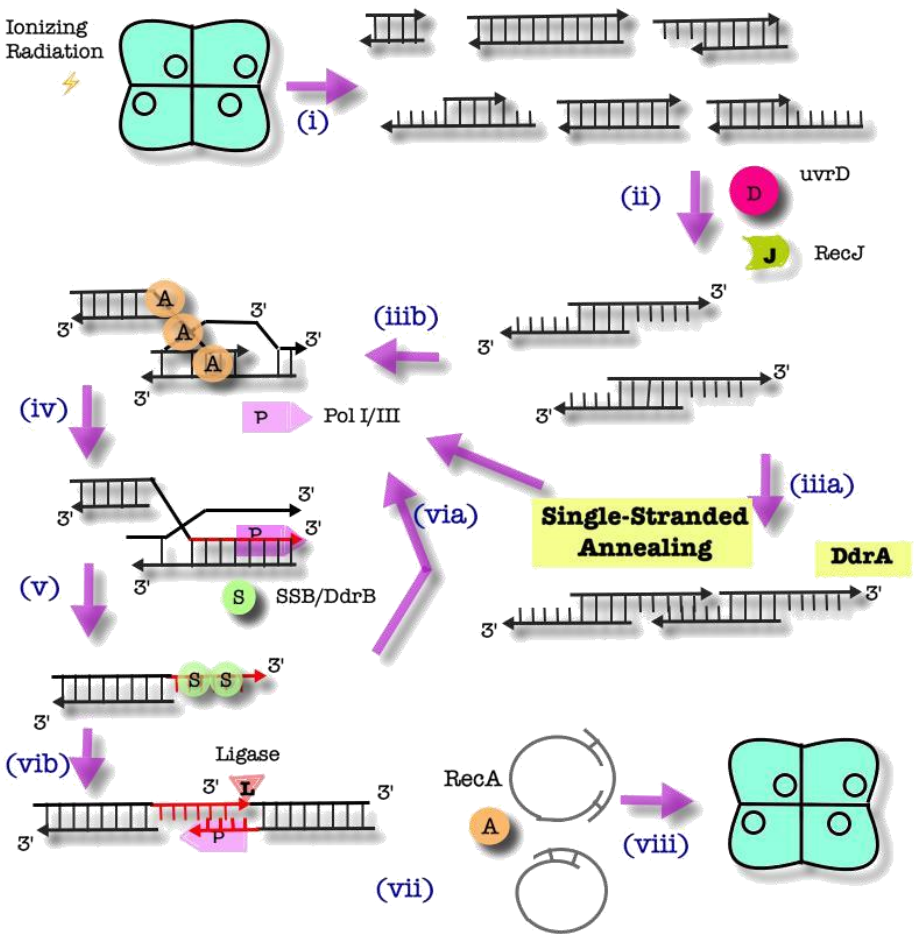


Figure 1.4: ESDSA Repair of Double-Strand Breaks in *D. Radiodurans*

i) Hundreds of double-stranded fragments are formed following a DNA damaging stimulus, such as ionizing radiation, which is pictured above.

ii) UvrD and RecJ then process the fragments to generate long 3' ssDNA strands.

iiia) Approximately one third of these fragments are then assembled into larger fragments via the mechanism of single-strand annealing.

Iiib) Fragments undergo RecA-mediated strand invasion.

iv) The invading 3' strand is then extended by a combination of pol I and pol III. The strand is then free to participate in either:

via) another round of strand invasion or

vib) association with a complementary fragment.

vii) Any gaps in the annealed fragments are filled and nicks are sealed by a ligase.

viii) The reassembled fragments are circularized by RecA-mediated HR.

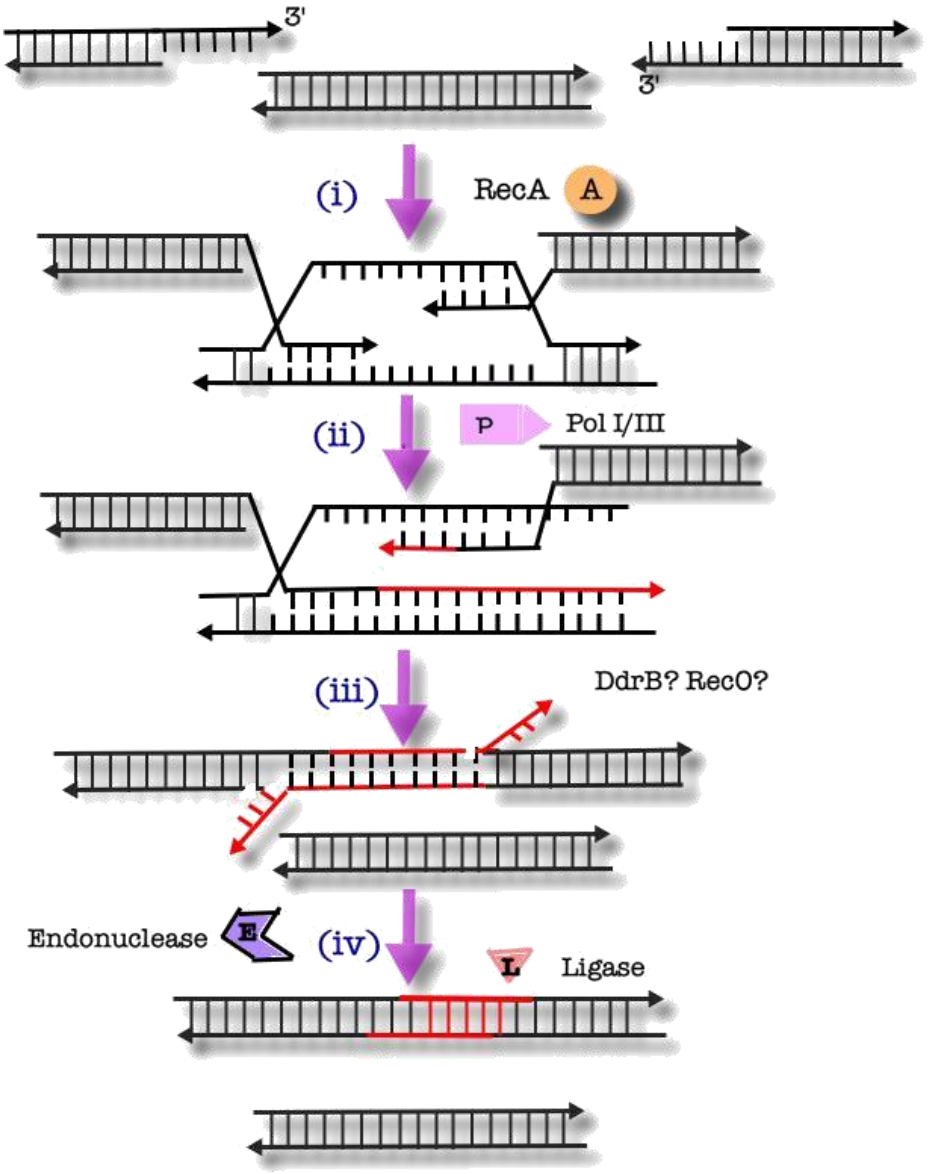


Figure 1.5: Synthesis Along a Bridging Element

i) Two DNA fragments, missing a segment of sequence between them, simultaneously invade a fragment containing the missing sequence.

ii) Both fragments are extended by a combination of pol I and pol III.

iii) The two fragments then disassociate from the bridging fragment and associate with each other.

iv) The extraneous sequence is excised by endonucleases, while the nicks are sealed by DNA ligase.

1.4.4 Single-Strand Annealing

DNA fragment assembly has been shown to occur in *D. radiodurans* in RecA-knockouts, suggesting that the process can occur independently of RecA and therefore, ESDSA (Daly & Minton, 1996; Slade *et al*., 2009). Following an ionizing radiation dose of 10 kGy, one third of all double-strand breaks are rejoined prior to RecA-mediated repair. Using this process, larger, partially repaired fragments are formed that may be better suited for subsequent repair by ESDSA (Daly & Minton, 1996). The RecA-independent repair occurs via a single-strand annealing (SSA) mechanism, similar to the one observed in *E. coli* (Daly & Minton, 1996; Kowalczykowski *et al*., 1994). In this model (Figure 1.4, iiia), a 3' ssDNA end generated by uvrD and RecJ is annealed to a complementary 3' ssDNA fragment from a separate genomic copy that was processed in the same manner. Any 5' flaps that remain after annealing are degraded and remaining gaps are filled in by DNA polymerase (Daly & Minton, 1996).

Work reported in this thesis demonstrates that the DNA damage response A (DdrA) protein possesses a novel ssDNA annealing activity that is required for DNA damage resistance in *D. radiodurans*.

1.4.5 Novel *Deinococcal* Proteins Involved in Repair

A full understanding of the proteins required for ESDSA and SSA is currently lacking. Only a few proteins have functionally assigned roles and most of these (e.g. RecA, Pol I, Pol III, RecFOR, uvrD, SSB and DNA ligase) are based on

analogy to what is known from their study in other systems (Slade *et al*., 2009). In addition to DdrA, two proteins (PprA and DdrB) have been shown to be critical for repair. All three proteins appear unique to the bacterium and have been identified in the genomes of all strains of *Deinococcus* that have been analyzed to date. DdrA, PprA and DdrB are upregulated following exposure to ionizing radiation, suggesting a role in repair. Furthermore, cells lacking DdrA, PprA or DdrB all exhibit radiosensitivity. A combined knockout of DdrA and DdrB results in greater radiosensitivity than knocking out either protein in isolation. Likewise, knocking out DdrA or DdrB together with RecA yielded greater radiosensitivity than knocking out RecA alone. Together, these findings suggest that DdrA and DdrB are epistatic to one another and to RecA. One possible interpretation of the data is that, similar to DdrA, DdrB may also enhance ssDNA annealing. If this hypothesis were correct, and the proteins were indeed responsible for performing redundant functions, it would explain why knocking out both proteins together leads to a greater reduction in radioresistance than knocking out either protein alone (Tanaka *et al*., 2004). Recent work in the Junop lab has demonstrated ssDNA annealing activity in DdrB (Sugiman-Marangos *et al*., 2016), adding strength to the idea that DdrA and B may share similar function in SSA repair.

In contrast, knocking out PprA and RecA together results in the same level of radiosensitivity as knocking out RecA alone, suggesting that PprA and RecA act in the same pathway. Knocking out RecA already eliminates the pathway that PprA is involved in and the elimination of PprA function therefore results in no further decrease in radioresistance.

1.5 DNA Damage Response A (DdrA)

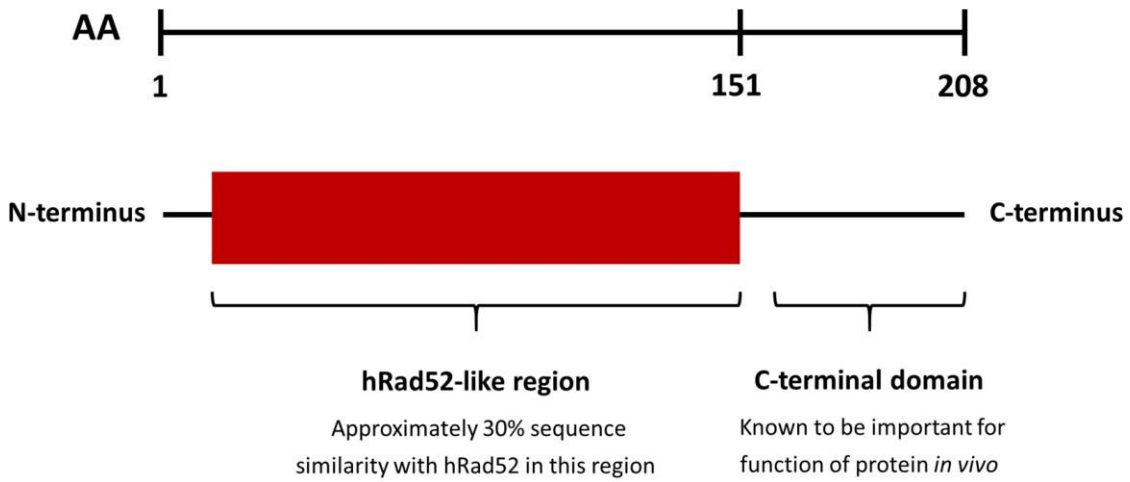
1.5.1 Previous Research

When work on this thesis began, no evidence of the ability of DdrA to enhance ssDNA annealing had been reported and no crystal structure of either the protein alone or the protein in complex with DNA had been determined. Nevertheless, the importance of DdrA for DNA damage resistance in *Deinococcus* had been clearly demonstrated. As mentioned in Section 1.4.5, DdrA was shown to be upregulated 23-fold following exposure to a sub-lethal dose of IR (3 kGy) (Harris *et al*., 2004). As well, cells lacking DdrA were found to be highly sensitive to IR and MMC (Harris *et al*., 2004). Work reported by Harris *et al*. (2004) further indicated that DdrA is incapable of binding dsDNA unless a 3' ssDNA extension is present. Providing the protein with dsDNA containing a 5' extension resulted in no significant interaction. Together, these findings suggest that the preferred DNA binding interaction of DdrA is with ssDNA containing a free 3' end. DdrA was also reported to lack ATPase, ssDNA annealing, helicase and recombinase activity (Harris *et al*., 2004).

In an effort to further define DdrA structure-function relationships, limited proteolysis was used to probe domain structure. These *in vitro* studies demonstrated that the N-terminal 157 residues of DdrA form a stable domain with full ssDNA binding activity. Deletion of the C-terminal protease sensitive region (residues 158-208) resulted in a partial loss of binding preference for ssDNA containing 3' versus 5' ends. Together, these findings suggested that the first 157

residues of DdrA form a stable, functional core domain. Surprisingly, when a gene encoding DdrA$^{1-157}$ was expressed *in vivo*, cells were as sensitive to IR as the knockout, indicating that the C-terminal region of DdrA plays an important role *in vivo* that is in addition to its interaction with ssDNA (Figure 1.6). The C-terminus of DdrA may serve a regulatory role or be required for interaction with other proteins and/or DNA structures. It is also possible that the C-terminal region may be involved in an activity that remains to be characterized (Harris *et al*., 2008). Determining a high-resolution structure, especially in complex with DNA, would offer mechanistic insight that might help address these and other questions surrounding DdrA function.

Although a crystal structure of DdrA has not been determined, a low-resolution electron microscopy (negative stain) structure was reported for the N-terminal domain (residues 1-160) of DdrA from *D. deserti* (Gutsche *et al*., 2008). The final reconstruction was determined to 23 Å and revealed a surprisingly complex quaternary structure. DdrA assembled into a 7-subunit heptameric ring that further self-associated into a trimer of ring structures yielding a final complex with 21 DdrA subunits (Gutsche *et al*., 2008). Since the interaction surface observed between ring structures was relatively small, it was suggested that DdrA would most likely exist as a heptameric ring in its biologically relevant state. Unfortunately, the low-resolution precluded further insight into structure-function relationships of DdrA.



1-157 (*D. radiodurans*): "stable, functional core" (Harris *et al*., 2008)
1-160 (*D. deserti*): negative stain EM structure (Gutsche *et al*., 2008)
1-160 (*D. geothermalis*): X-ray diffraction to 2.4 Å (Chapter 3 of this thesis)
1-161 (D. radiodurans): ssDNA annealing (Chapter 2 of this thesis)

Figure 1.6: Domain Organization of DdrA. The first 151 residues of DdrA constitute a Rad52-like domain, whereby approximately 30% of the sequence is shared with Rad52. The thread-based homology modelling, which yielded these results, is detailed in Chapter 3. The remainder of the protein constitutes a C-terminal domain, which is important for radioresistance *in vivo* (Harris *et al*., 2008). The first 157 residues have been shown to be sufficient in forming a stable, functional core as DdrA$^{1-157}$ (D. *radiodurans*) displayed a nearly identical biochemical profile *in vitro* as the full-length protein. DdrA$^{1-160}$ (*D. deserti*) was useful in determining a low-resolution heptameric structure of the protein by negative stain EM. DdrA$^{1-160}$ (*D. geothermalis*) was useful in obtaining a crystal that diffracted to 2.4 Å (see Section 3.4.4). The work reported in Chapter 2 has shown that DdrA$^{1-161}$ (*D. radiodurans*) possesses robust strand annealing activity.

1.5.2 Current Research Outline

Given the importance of DdrA for DNA damage tolerance in *Deinococcus,* a primary objective of this thesis was to further characterize the structure-function relationships of DdrA using a combination of biophysical and cell-based techniques. The primary goal for structural characterization was to determine an X-ray crystal structure of DdrA in complex with DNA. Such information would not only inform on potential mechanisms for the protein, but also provide a framework for further hypothesis-based studies. In Chapter 3, we outline steps taken to obtain optimal constructs of DdrA suitable for crystallization. In addition, conditions for the successful crystallization of DdrA are reported.

Although DdrA appears to be unique to *Deinococcus*, it does share some weak sequence similarity with a domain of eukaryotic Rad52 (residues 1-209) responsible for binding ssDNA and enhancing strand annealing (Singleton *et al*., 2002). We therefore sought to determine if DdrA might also be capable of annealing ssDNA. In Chapter 2, we report the identification of robust annealing activity within the first 161 residues of DdrA, comparable to human Rad52. This finding is particularly significant as it contradicts a previous report that suggested DdrA is incapable of annealing ssDNA (Harris *et al*., 2004). Importantly, we further establish, through mutational studies, that this novel annealing activity is required for the ability of DdrA to function in its role of promoting extreme DNA damage tolerance in *D. radiodurans*.

Chapter 2: Functional Characterization of DdrA

2.1 Abstract

Deinococcus radiodurans has several unique proteins required for its extraordinary resistance toward a wide range of DNA damaging stimuli. DdrA represents one of these proteins and is thought to be directly involved in DNA double-strand break repair. Although DdrA shares weak sequence similarity with human Rad52 (<10% identity), no prokaryotic Rad52-like homologue has been shown to possess ssDNA annealing activity, suggesting that DdrA may perform a different role in *Deinococcus*. To further characterize DdrA, we tested the possibility that DdrA might function as a ssDNA annealing protein. Contrary to prior reports, DdrA was found to possess robust ssDNA annealing activity. This activity was localized to an N-terminal domain (residues 1-161) that appears to be partially regulated by elements in the less structured C-terminal region (residues 161-208). Two residues (K22 and K105) necessary for ssDNA annealing were identified and used to demonstrate a requirement for DdrA annealing activity in *Deinococcus* following exposure to extreme levels of DNA damage. Taken together, this work not only suggests that DdrA functions as a Rad52 homologue for annealing of ssDNA, but also represents the first demonstration that any prokaryote contains both a functional and structural Rad52 homologue.

2.2 Introduction

Double-strand breaks are perhaps the most lethal form of damage sustained by DNA. In lower organisms, a single DSB may prove to be lethal if left unrepaired. In higher, multicellular organisms, erroneous repair of DSB’s is

associated with mutagenic events that can lead to a variety of deleterious outcomes, including cancer (Pardo *et al*., 2009).

The bacteria of the genus *Deinococcus* possess a remarkable capacity to recover from exposure to high levels of DNA damage. The model organism *Deinococcus radiodurans* is capable of surviving approximately 15 kGy of gamma radiation, which effectively shatters the genome into hundreds of fragments, each approximately 20 to 30 kb in length (Zahradka *et al*., 2006). Remarkably, the complete genome of the bacterium is faithfully reassembled from these fragments in a matter of hours. In contrast, humans are several thousand times more sensitive to gamma radiation with a lethal dose in the range of 2-10 Gy (Mihandoost *et al*., 2014).

Since the discovery of *D. radiodurans* in 1956 (Anderson *et al*., 1956), numerous hypotheses have been proposed in an effort to explain this remarkable degree of survival. It has been demonstrated that *D. radiodurans* accumulates DNA damage to the same extent as radiosensitive organisms, such as *E. coli,* meaning that the bacterium does not protect its genome prophylactically. Instead, radioresistance is thought to result from a combination of 1) protection of repair proteins by elevated intracellular concentrations of manganese and 2) robust repair pathways reliant on a collection of seemingly unique proteins (Zahradka *et al*., 2006; Tanaka *et al*., 2004; Slade *et al*., 2009; Makarova *et al*., 2007).

The restoration of the *Deinococcal* genome in response to severe irradiation occurs via a two-stage process. The first stage, ESDSA, produces 3’ ssDNA

extensions greater than 20 kb in length, which result from successive rounds of strand invasion of homologous fragments followed by Pol I and Pol III-mediated strand extension (Zahradka *et al*., 2006; Slade *et al*., 2009). In the second stage of repair, complete circular chromosomes are generated by the joining of long linear DNA molecules via RecA-dependent homologous recombination. Interestingly, fragment assembly has also been observed in *Deinococcal* cells lacking RecA function, indicating that the process is capable of occurring independently of ESDSA (Daly & Minton, 1996; Slade *et al*., 2009). In fact, approximately one third of all double-strand breaks resulting from a dose of ionizing radiation of 10 kGy are repaired in a RecA-independent manner. It is thought that larger DNA fragments, generated by RecA-independent repair, serve as more ideal substrates for ESDSA (Daly & Minton, 1996). This RecA-independent process is believed to occur through a single-strand annealing mechanism (Daly & Minton, 1996; Kowalczykowski *et al*., 1994), similar to the one observed in *E. coli*.

Despite knowing that ESDSA and SSA are essential for extreme DNA damage tolerance, a complete list of proteins responsible for carrying out various functions in these pathways is not yet available. It is already clear, however, that some functions are fulfilled by ‘house-keeping’ repair proteins (i.e. RecFOR) while other functions are completed by proteins (such as DdrA, DdrB and PprA) that are only needed for repair of extreme amounts of DNA damage. This latter class of proteins is of particular interest since they are likely to be directly responsible for mediating the remarkable DNA damage resistance of *Deinococcus.*

DdrA has been shown to be essential for extreme DNA damage resistance, but its mechanism of action has remained elusive. Although DdrA shares weak sequence similarity (<10% identity) with human Rad52, it was thought that DdrA does not function as a Rad52 homologue for annealing of ssDNA. This thinking was based on the fact that very few Rad52-like homologues have been identified in prokaryotes and that there already exists a well characterized functional homologue (RecO, not structurally related to Rad52) of Rad52 sufficient for mediating strand annealing during RecA-directed repair. Nevertheless, no firm data have been reported to conclusively rule out the possibility that DdrA functions as a ssDNA annealing factor uniquely required for extreme DNA damage tolerance.

In this chapter, we demonstrate robust ssDNA annealing activity in DdrA. This represents the first report of such activity for any Rad52-like protein in any prokaryote. Annealing activity was further localized to an N-terminal (residues 1-161) domain that appears to be partially regulated by elements in the less structured C-terminal region (residues 161-208). Furthermore, we identify amino acid residues of DdrA (K22 and K105) required for binding and annealing of ssDNA. Most significantly, we present evidence suggesting that DdrA annealing activity is essential for *Deinococcus* to achieve extreme levels of DNA damage resistance.

2.3 Materials & Methods

2.3.1 Protein Preparation

Gateway Cloning

The gene encoding full-length DdrA (*D. radiodurans*) was synthesized (GenScript) and codon optimized for expression in *E. coli*. DdrA was Gateway cloned using the pUC57 entry vector obtained from GenScript into a pDEST527 expression vector encoding an N-terminal polyhistidine-tag and tobacco etch virus (TEV) protease cleavage site. To determine whether the C-terminus is required *in vivo* due to reasons related to ssDNA annealing, we also prepared $DdrA^{1-161}$ (*D. radiodurans*). This protein was similarly Gateway cloned but placed into a pDEST14 expression vector. Unlike the pDEST527 expression vector, the pDEST14 expression vector does not add any fusions. Instead, a C-terminal polyhistidine-tag was engineered during gene synthesis. The integrity of the two expression vectors was verified by Sanger sequencing. The bacterial expression vector (pSF2285) encoding full-length human Rad52 (with an N-terminal His_{6x} fusion and SUMO protease cleavage site) was obtained as a kind gift from Dr. Mauro Modesti (Centre de Recherche en Cancérologie de Marseille).

Protein Expression and Purification

All proteins were expressed in *E. coli* BL21 (DE3)-T1R cells (Invitrogen™). Cultures were grown in lysogeny broth (LB) supplemented with ampicillin (100 µg/mL) at 37°C to an optical density 600 (OD_{600}) of approximately 0.5 and induced

with 1 mM of isopropyl β-D-thiogalactopyranoside (IPTG) for 16 hours at 16°C. Cells were harvested by centrifugation at 6000 x g for 15 min at 4°C. Cell pellets were resuspended in lysis buffer (1 M NaCl, 20 mM Tris-HCl pH 8.0 and 5 mM imidazole) using 10 mL buffer per gram of cell pellet and lysed by 4 sequential French press passages at 10,000 psi. Following clarification by centrifugation at 48,384 x g for 40 min, soluble lysate was loaded onto a Ni-charged HisTrap™ Fast Flow (FF) 5 mL column (GE Healthcare) at 1 mL/min using an ÄKTA Fast Protein Liquid Chromatography (FPLC) system. The column was washed with 15 column volumes each of buffer (1 M NaCl and 20 mM Tris-HCl pH 8.0) containing increasing amounts of imidazole (0, 5, 15, 30, 45 and 60 mM imidazole) prior to elution (600 mM imidazole). Eluted protein was buffer exchanged using a HiPrep 16/10 desalting column (GE Healthcare), equilibrated with TEV protease buffer, composed of 150 mM NaCl, 50 mM Tris-HCl pH 8.0, 5 mM dithiothreitol (DTT) and 0.5 mM ethylenediaminetetraacetic acid (EDTA).

While the C-terminal polyhistidine-tag of $DdrA^{1-161}$ was designed to be uncleavable, full-length DdrA and Rad52 were digested using TEV and SUMO proteases, respectively. Small-scale assays were first performed to determine optimal conditions for cleavage. Large-scale cleavage reactions were performed with a 10:1 ratio of fusion protein to protease. Following digestion with TEV or SUMO protease, samples were exchanged into lysis buffer (1 M NaCl, 20 mM Tris-HCl pH 8.0 and 5 mM imidazole) and passed over a 5 mL Ni-charged HisTrap™ column to isolate digested protein. Proteins were concentrated by ultrafiltration (10k molecular weight cut-off [MWCO], VivaSpin), aliquoted and stored at -80°C.

2.3.2 Analysis of Quaternary Structure

To assess monodispersity of purified DdrA and estimate quaternary structure, size-exclusion chromatography (SEC) was performed. DdrA (10 mg/mL) was resolved on a HiLoad™ 16/60 Superdex™ 200 prep grade column (GE Healthcare) using an ÄKTA Pure system (GE Healthcare) housed at 10°C. The column was equilibrated and run with buffer containing 20 mM Tris-HCl pH 8.0, 1 M NaCl and 15% glycerol (v/v). Molecular weight standards were run under the same conditions to calibrate the column for size estimation of DdrA.

Size-exclusion chromatography with multi-angle light scattering (SEC-MALS) was also performed under the same conditions with the SEC column connected in-line to a Dawn HELEOS II MALS detector equipped with a 662 nm laser source and an Optilab T-rEX differential refractometer equipped with a 658 nm LED source (Wyatt Technology, Santa Barbara, CA, USA). Molecular weights were calculated by Zimm plot analysis using ASTRA software (v6.1.5.22; Wyatt Technology).

2.3.3 ssDNA Binding Assessment

ssDNA Synthesis

As the 3’ end of ssDNA has been shown to be important for ssDNA-DdrA interactions, oligonucleotides were labelled (BioBasic) at the 5’ end with an Alexa Fluor488 for the purpose of visualization. Since SEC analysis suggested that DdrA adopts a heptameric arrangement, DNA sizes were synthesized in multiples of 7

(i.e. 14, 21, 28, 35, 42 and 49 nt oligomers) so that each monomer of the protein could interact with a whole number of nucleotides. To minimize the potential for secondary structure formation, oligonucleotides were designed entirely of either A or T bases.

Electrophoretic Mobility Shift Assays (EMSA's)

DdrA ssDNA binding activity was determined using electrophoretic mobility shift analysis. ssDNA oligomer (20 nM) was mixed with increasing amounts of DdrA (0.16, 0.32, 0.64, 1.25, 2.5, 5, 7.5 and 10 μM) to a final volume of 20 μL in buffer containing 150 mM NaCl, 10% (w/v) glycerol, 40 mM Tris-HCl pH 8.0, 0.1 mM EDTA and 1 mM DTT. Following incubation at 30°C for 30 min, the extent of DNA binding was determined by electrophoresis using native Tris/Borate/EDTA (TBE) 20% polyacrylamide gels. Gels were imaged using a ChemiDoc™ Gel Imaging System and DNA bands quantified using ImageJ software.

2.3.4 ssDNA Annealing Assay

ssDNA Oligomers

Once DdrA had been confirmed to bind ssDNA, annealing capacity was evaluated in comparison to Rad52. Two perfectly complementary 48 nt oligomers served as substrates for the assay. The forward oligomer, 5′-GCAATTAAGCTCT AAGCCATCCGCAAAAATGACCTCTTATCAAAAGGA-3', was synthesized in two forms: a fluorescently labelled variety, from this point forward, referred to as "oligo 1-F" and an unlabelled form, referred to as "oligo 1 -U". The complementary reverse

oligomer (oligo 2) was unlabelled. Oligo 1-F was labelled at the 5' end to avoid disruption with DdrA interaction. All oligomers were purchased from BioBasic.

ssDNA Annealing

To assess ssDNA annealing capacity, oligo 1-F was pre-incubated in the absence or presence of protein (DdrA, $DdrA^{1-161}$ and hRAD52) at 10°C for 1 h in a 20 µL volume of annealing buffer (30 mM Tris-acetate pH 7.5, 5 mM magnesium acetate and 1 mM DTT). Oligo 2 was likewise pre-incubated separately under the same conditions. Annealing reactions were initiated by combining pre-incubated mixtures of oligo 2 and oligo 1-F. Final reactions (40 µL total) contained 200 nM of each oligonucleotide and 1 µM of protein. Reactions were incubated at 10°C and stopped at varying time points (i.e. 0, 1, 2, 5, 10, 15, 20 or 25 mins) by the addition of a 40-fold excess of oligo 1-U. No protein and protein conditions were run in parallel for each experiment to minimize error between replica experiments. Each annealing assay (for a given protein) was repeated a minimum of three times.

The extent of ssDNA annealing was determined by resolving ssDNA and dsDNA on 10% polyacrylamide gels. DNA species were visualized using a ChemiDoc™ Imaging System. The relative intensities of the bands representing the annealed and unannealed species were quantified using Bio-Rad's Image Lab software. This information was then used to determine the percentage of ssDNA annealed as a function of time.

2.3.5 *In Vivo* Analysis of ssDNA Annealing by DdrA

Mutant Design

To test the importance of DdrA ssDNA annealing *in vivo*, a mutant of DdrA lacking annealing activity needed to be identified. A multiple sequence alignment of all annotated DdrA homologues was generated to identify absolutely conserved, positively charged residues that might mediate ssDNA annealing. By including hRad52 in the alignment, the number of absolutely conserved positively charged residues was reduced to 6. To further limit the number of residues for mutagenesis and functional testing, a thread-based model of DdrA was generated using **P**rotein **H**omology/Analog**y** **R**ecognition **E**ngine Version **2**.0 (Phyre2). This model was used to determine which conserved residues of DdrA might be structurally analogous to residues in hRad52 (R55 and K152) known to be essential for ssDNA annealing (Honda *et al*., 2011). These residues then served as targets for mutagenesis.

Survival Assay

Once DdrA$^{K22/K105}$ was confirmed to lack annealing activity *in vitro*, an *in vivo* survival assay was performed to assess the importance of DdrA ssDNA annealing in response to DNA damage. Wild-type *D. radiodurans* cells (positive control) and ΔDdrA *D. radiodurans* cells (negative control) were transformed with empty vectors to allow for chloramphenicol selection. ΔDdrA cells were also transformed with a vector encoding DdrA$^{K22/K105}$. One microgram of the appropriate DNA was added to 100 µL of cells in a 50 mL sterile tube. The solution was gently mixed and placed

on ice for 10 minutes followed by incubation at 32°C for 30 minutes, accompanied by gentle shaking. One millilitre of sterile TGY media was added to the transformation mixture followed by incubation at 32°C with vigorous shaking. After 18 h, 100 µL of the transformation mixture were plated onto TGY agar plates supplemented with 3 µg/mL of chloramphenicol. Surviving colonies were picked and cultures were grown to supersaturation ($OD_{600} \approx 2$) in TGY media at 32°C. The cultures were then serially diluted, first to an OD_{600} of 0.5 and then to an OD_{600} of 0.1. In addition to $OD_{600} = 0.1$, cultures with readings of 0.1×5^{-1}, 0.1×5^{-2}, 0.1×5^{-3} and 0.1×5^{-4} were also prepared. 10 µL of each dilution was then spotted onto TGY agar plates containing different concentrations of mitomycin C (0, 25, 50, 75, 100 and 125 ng/mL). Colony survival was monitored over the course of several days. Data from all five MMC dilutions were averaged to generate survival curves (as seen in Figure 2.4.7).

2.4 Results

2.4.1 Expression and Purification of DdrA for *in Vitro* Functional Studies

In order to test the ssDNA binding and annealing activities of DdrA *in vitro*, we first needed to express and purify soluble DdrA. Since the Rad52-like domain of DdrA (residues 1-161) is not functional for radioresistance in *Deinococcus* (Harris *et al*., 2008), we chose to test DNA binding and annealing activities of both full-length and C-terminally truncated ($DdrA^{1-161}$) protein. Domain boundaries for truncated DdrA were chosen based on comparison of secondary structure

prediction and sequence conservation with hRad52 (see Figure 3.1.1). Each construct was generated with a His_6 fusion to aid in purification.

All proteins were expressed in BL21(DE3)-T1R cells at 16 ºC for 16 h. Each protein was expressed at high levels and could be isolated in soluble form. Proteins were purified using immobilized metal affinity chromatography (IMAC) (see Section 2.3.1). Briefly, protein was captured from soluble lysate using a 5 mL nickel-charged IMAC column. Loosely bound impurities were removed by successive washes with buffer containing increasing concentrations of imidazole prior to elution. Results from a representative purification are illustrated in Figure 2.1. For full-length DdrA and hRad52, the N-terminal His tag was removed by treatment with TEV and SUMO proteases, respectively. Samples were then further purified by IMAC and SEC. At this stage, proteins were estimated to be greater than 95% pure. Final yields ranged between 2-5 mg of pure protein per L of cell culture.

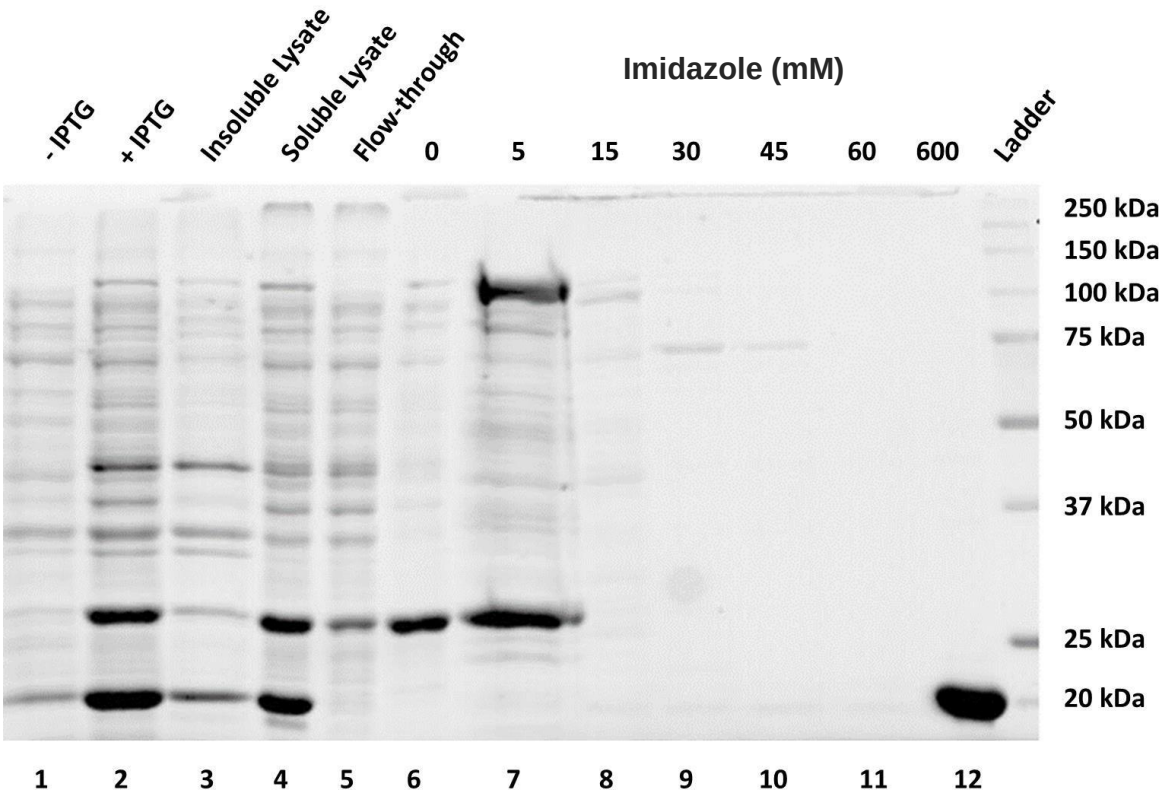


Figure 2.1: IMAC Purification of DdrA$^{1\text{-}161}$. Lanes 1 and 2 depict the proteins present prior to and following the addition of IPTG, respectively. Lanes 3 and 4 are protein in the soluble and insoluble fractions (respectively) following centrifugation. Lane 5, proteins that failed to bind the nickel column. Lanes 6-11 depict elution from step-washes at increasing amounts of imidazole. Lane 12, final IMAC elution fraction. DdrA$^{1\text{-}161}$ has a theoretical mass of 18.6 kDa, corresponding to the apparent molecular weight of eluted protein.

2.4.2 Quaternary Structure Findings

In order to assess protein quality (i.e. monodispersity) and quaternary structure, size exclusion chromatography was performed (Figure 2.2). Elution retention volumes were used to estimate apparent molecular weight based on calibration performed with protein molecular weight standards. As shown in Figure 2.2, DdrA$^{1-161}$ eluted as a monodisperse species at a volume consistent with an apparent molecular weight of ~130 kDa, suggesting that DdrA$^{1-161}$ adopts a heptameric quaternary structure. Full-length DdrA was found to elute over a broader volume range close to the void volume, suggesting that full-length DdrA adopts a multitude of very large (>600 kDa) protein assemblies. Taken together, these findings suggest that the C-terminal region of DdrA may mediate protein-protein interactions required for higher-order assembly of multiple heptameric rings, similar to what has been observed for DdrA using negative stain EM (Gutsche *et al*., 2008). Although EM studies were able to reconstruct a trimeric arrangement of DdrA heptamers (i.e. with 21 subunits), it is likely that under varied conditions of buffer and protein concentration, other higher-order assemblies might also exist in solution. The functional significance of such assemblies, however, remains to be determined.

The quaternary structure of the full-length protein was further characterized by SEC-MALS. In this analysis, protein samples were run using a HiPrep™ 16/60 Sephacryl™ S-300 High-Resolution column, which is capable of resolving proteins up to 1.5 MDa in size. However, the same phenomenon was observed, whereby

the protein eluted too close to the void volume of the column, causing the light scattering intensity to be immensely overestimated. The results of the SEC-MALS were thus inconclusive and suggestive of a very large quaternary structure in the MDa range.

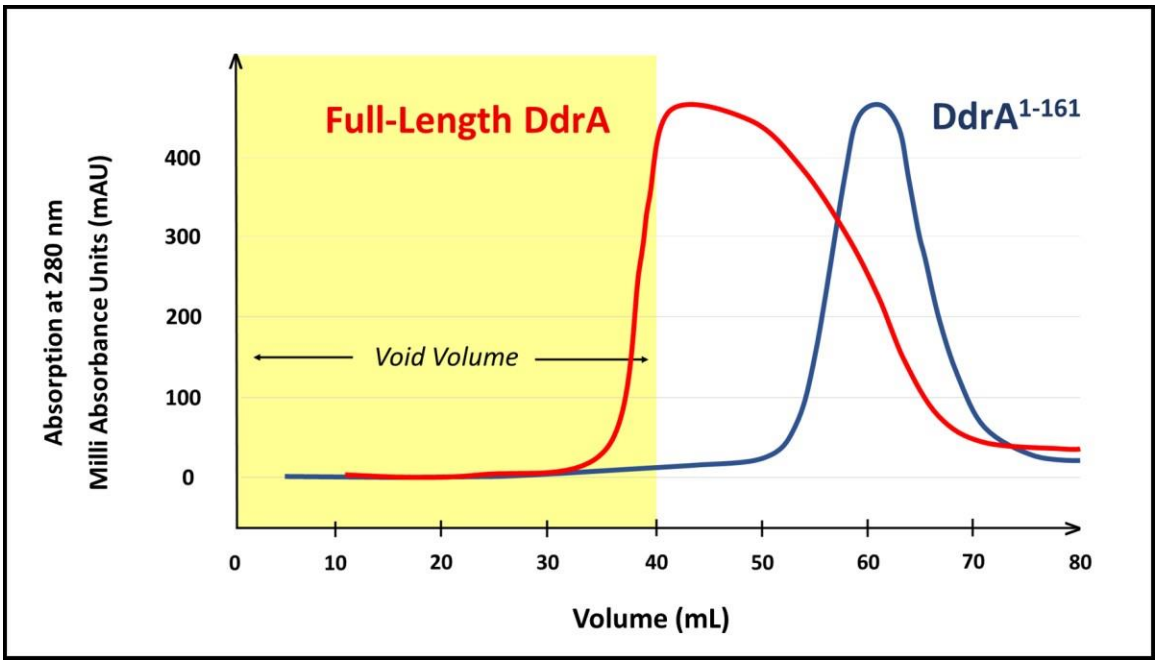


Figure 2.2: Analysis of Quaternary Structure of DdrA by SEC. Full-length DdrA eluted close to the void volume of the HiLoad™ 16/60 Superdex™ 200 prep grade column, suggesting a very large protein assembly. The column was calibrated using a series of protein standards (Ferretin [440 kDa], aldolase [158 kDa], conalbumin [75 kDa] and ovalbumin [43 kDa] eluted at 45 mL, 53 mL, 70 mL and 80 mL). Based on this calibration, DdrA$^{1-161}$ eluted at a volume consistent with an apparent molecular weight of ~130 kDa, suggesting that DdrA$^{1-161}$ adopts a heptameric conformation. The sample of truncated DdrA appears to be very monodisperse compared to the sample of full-length DdrA.

2.4.3 Characterization of DdrA ssDNA Binding

Before testing whether DdrA is capable of ssDNA annealing, we first sought to characterize the effect of ssDNA length on interaction with DdrA. Both full-length DdrA and DdrA$^{1-161}$ were incubated with oligonucleotides that varied in length by multiples of seven so that each monomer of the heptamer could interact with a whole number of nucleotides. An example of an EMSA that was generated can be seen on the left-hand side of Figure 2.4.3. In summation, both full-length and truncated DdrA were able to bind all lengths of ssDNA tested (14 to 49 nt). However, affinity increased with ssDNA length up to 28 nt, at which point, it plateaued (Kd ~ 8 μM with 14 nt and ~ 0.5 μM with ≥ 28 nt). Taken together, these data suggest that DdrA interacts most optimally with DNA when 4 nucleotides are bound to each monomeric copy of the protein. These findings are consistent with a report, which demonstrated that each monomer of Rad52 binds 4 nt of ssDNA (Parsons *et al*., 2000). There were no appreciable differences between poly-A or poly-T binding as one would expect for a protein that is required to repair any DNA sequence. Similarly, there were no appreciable differences in the binding capacities of the full-length and truncated forms of the protein, suggesting that the C-terminal tail may not play a significant role in ssDNA binding *in vitro*.

2.4.4 Analysis of DdrA ssDNA Annealing

Once DdrA (full-length and truncated) was confirmed to bind ssDNA, annealing activity was evaluated in an effort to characterize the role of DdrA in *Deinococcal* DNA repair. Several methods have been reported in the literature for measuring ssDNA annealing, including methods based on Förster resonance energy transfer (FRET), 4′,6-diamidino-2-phenylindole (DAPI) binding and gel electrophoresis. We conducted our preliminary experiments using the DAPI method, which makes use of the fact that DAPI stain has an increased affinity for duplex DNA and only fluoresces strongly when bound to dsDNA. This method was, however, abandoned due to problems associated with DdrA intrinsic fluorescence and non-specific interaction with the dye. We therefore decided to use a gel-based approach in which the annealing of a fluorescently labelled oligonucleotide to an unlabelled complementary strand was measured via gel electrophoresis. Since the ssDNA annealing activity of human Rad52 has been well characterized, we chose to use it as a positive control. As expected, hRad52 was found to greatly enhance the rate of ssDNA annealing, validating our annealing assay (Figure 2.3.1).

Using this assay, we found that both full-length and truncated forms of DdrA were able to enhance ssDNA annealing (Figures 2.3.2 & Figure 2.3.3). As shown in Figure 2.3.4, $DdrA^{1-161}$ annealing efficiency was comparable to hRad52, and both were significantly more active than full-length DdrA. This is particularly interesting since $DdrA^{1-157}$ has previously been shown to lack the ability to confer

radioresistance *in vivo*. Taken together, these findings suggest that the reason a C-terminal truncation of DdrA failed to complement a DdrA knockout is not due to an inability to anneal ssDNA. Rather, it would appear that the C-terminal tail of DdrA is essential for repair *in vivo* in a manner that is independent of ssDNA annealing. Indeed, the analogous region of hRad52 has been shown to mediate essential interactions with other binding partners such as Rad51 (Kagawa *et al*., 2014).

In vitro, the presence of the C-terminal tail of DdrA appeared to reduce ssDNA annealing activity. This may reflect the fact that truncation of the C-terminal region alters quaternary structure from very large assemblies to a more simplified heptamer. Alternatively, it may point to a 'regulatory' role for the C-terminal region, perhaps limiting annealing to instances when the cell has undergone extensive amounts of DNA damage that cannot be repaired by RecO-mediated events.

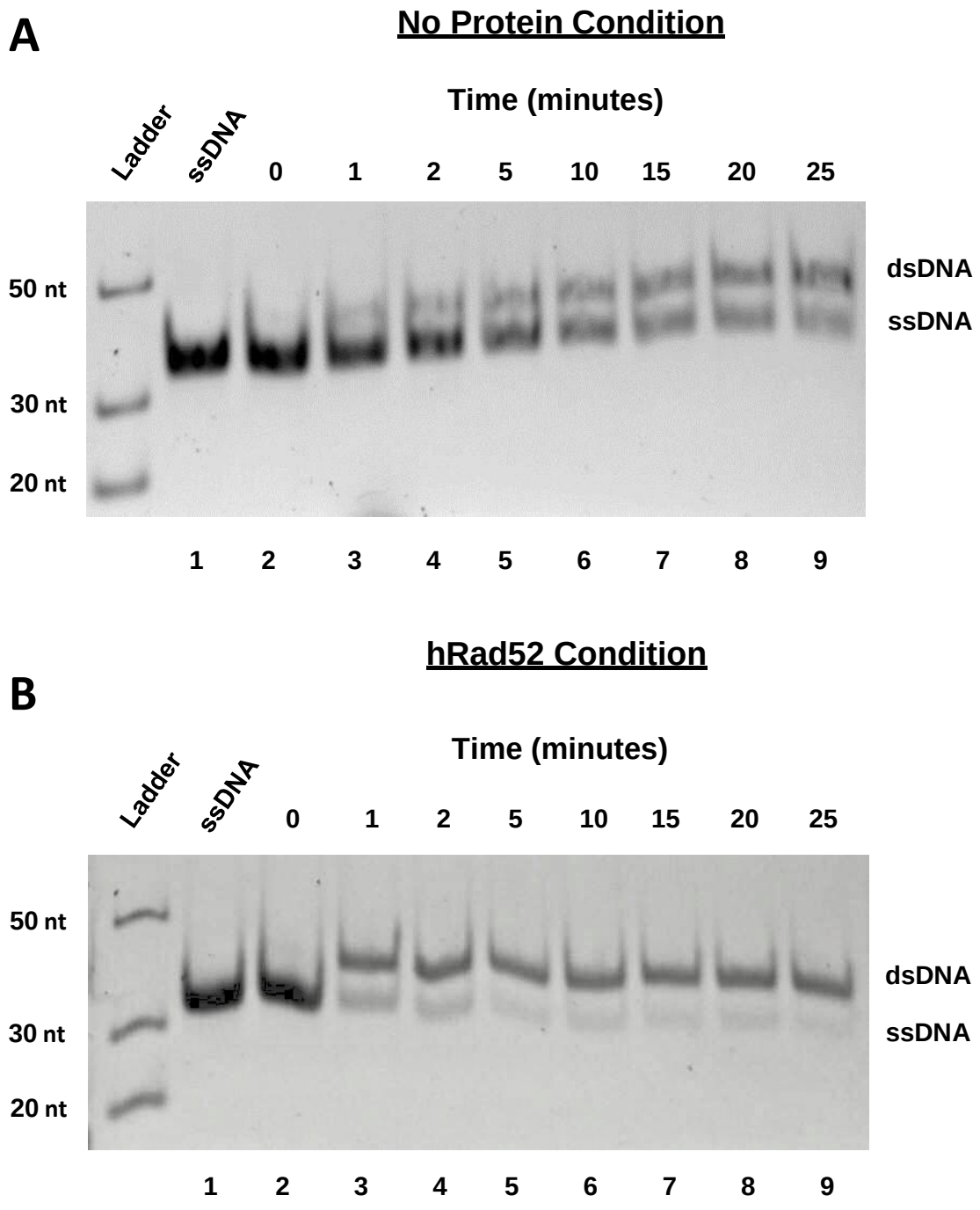


Figure 2.3.1: ssDNA Annealing Assessment of hRad52. ssDNA annealing is depicted in the absence (A) and presence of hRad52 (B). Lane 1 depicts ssDNA only, for comparison. The results suggest that Rad52 is capable of significantly enhancing strand annealing.

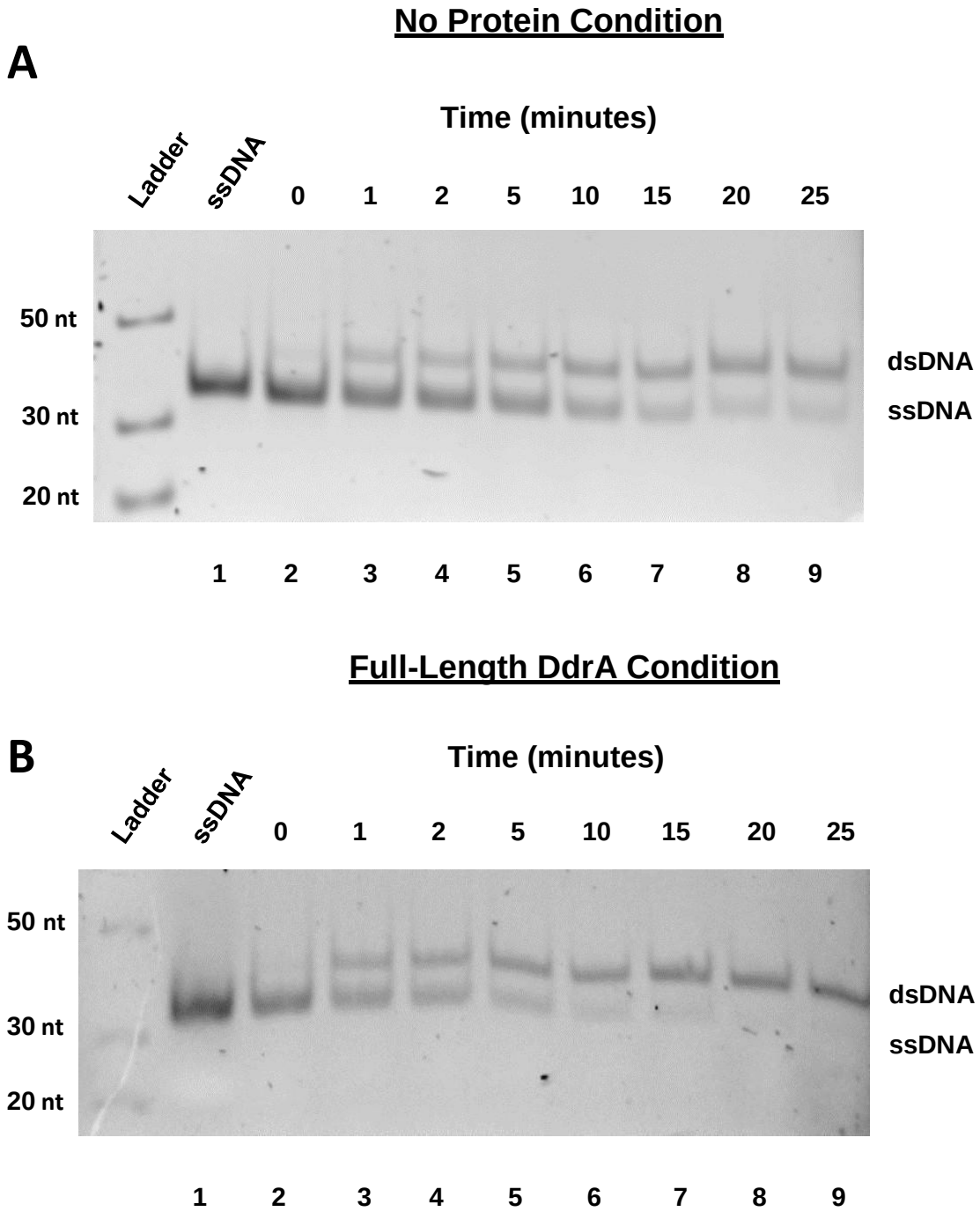


Figure 2.3.2: ssDNA Annealing Assessment of Full-Length DdrA. ssDNA annealing is shown in the absence of protein (A) and in the presence of full-length DdrA (B). The results indicate that full-length DdrA is capable of modest strand annealing enhancement.

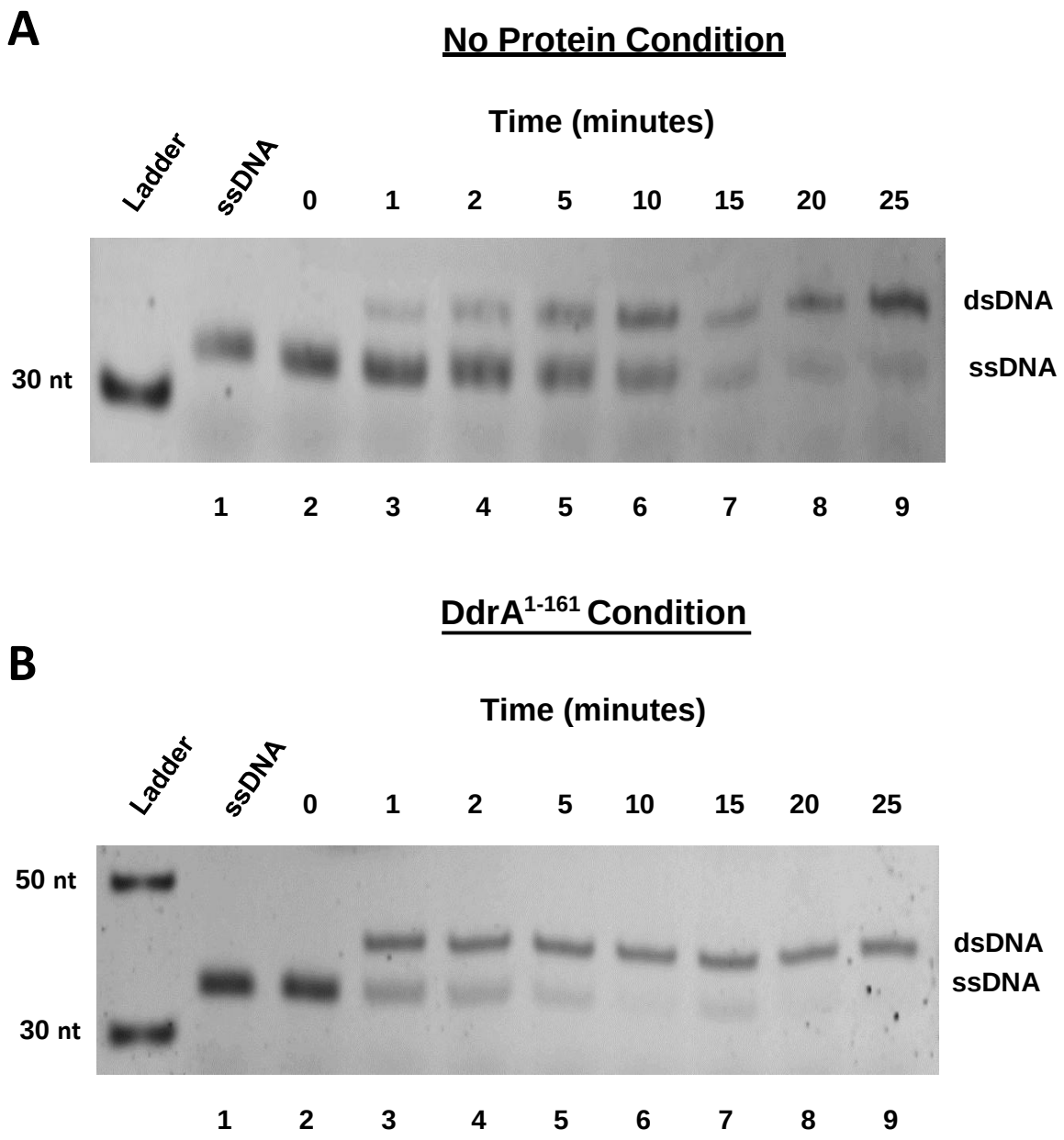


Figure 2.3.3: ssDNA Annealing Assessment of DdrA$^{1-161}$. ssDNA annealing is depicted both in the absence of protein (A) and under the influence of DdrA$^{1-161}$ (B). The results suggest that the truncated protein has a higher annealing activity than its full-length counterpart, closer to the annealing activity of Rad52.

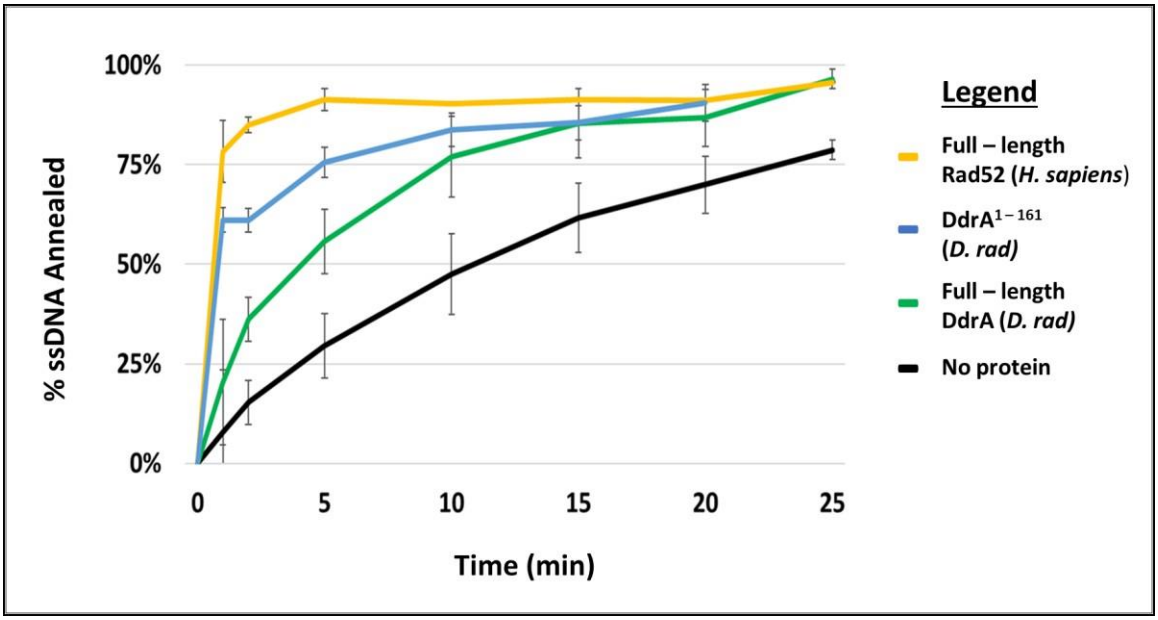


Figure 2.3.4: Summary of ssDNA Annealing Data. The percentage of ssDNA annealed, as a function of time, is shown. The black line represents the unassisted condition, whereby no protein was added. The ssDNA annealing process is enhanced when full-length DdrA (green) is added and even more so when DdrA$^{1-161}$ (blue) is introduced. The greatest annealing effect is observed following the addition of hRad52 (yellow), which was used as a positive control for the assay. Each data point represents the mean of three separate, independent trials, while the error bars are representative of the standard deviations.

2.4.5 Importance of DdrA ssDNA Annealing for DNA Damage Tolerance

In order to assess the biological relevance of ssDNA annealing by DdrA, we first needed to identify specific residues that mediate this process. We therefore compared DdrA proteins from seven different strains of *Deinococcus* with Rad52 proteins from *Thermus aquaticus* (*T. aquaticus*) and *Homo sapiens* (*H. sapiens*) using a multiple sequence alignment. As shown in Figure 2.4.1, the alignment indicated that there were six different basic residues that were absolutely conserved. In order to further define which of these residues might be involved in ssDNA annealing, we compared the predicted structure of DdrA to the known structure of Rad52 using thread-based homology modelling. As seen in Figure 2.4.2, residues K22 and K105 of DdrA were found to align well with residues R55 and K152 of hRad52. Importantly, R55 and K152 have been shown to be essential for mediating ssDNA annealing in Rad52 (Saotome *et al*., 2018). K22 and K105 of DdrA were therefore substituted to alanine residues in order to disrupt interaction with the negatively charged backbone of DNA.

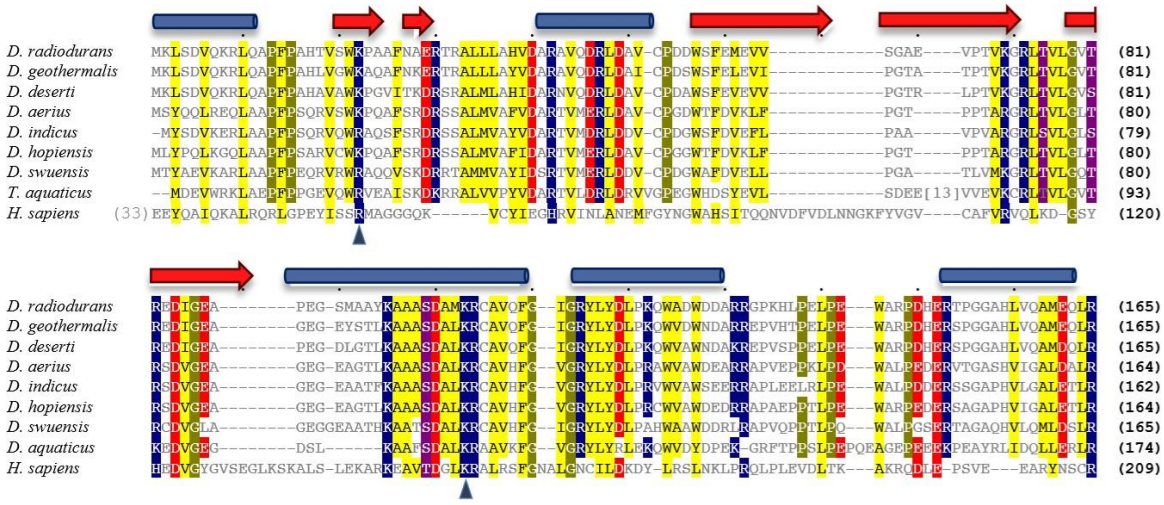


Figure 2.4.1: Mutant Design: Multiple Sequence Alignment. Amino acid comparison of DdrA proteins from seven different strains of *Deinococcus* and Rad52 proteins from *T. aquaticus* and *H. sapiens*. Instances where specialized amino acids, such as glycine and proline, are conserved are highlighted in gold, polar uncharged residues in purple, basic residues in blue, acidic residues in red and hydrophobic residues in yellow. The red arrows and blue cylinders represent beta strands and alpha helices of hRad52, respectively. Blue arrows indicate the position of two conserved lysine residues that were substituted to alanine residues for functional studies.

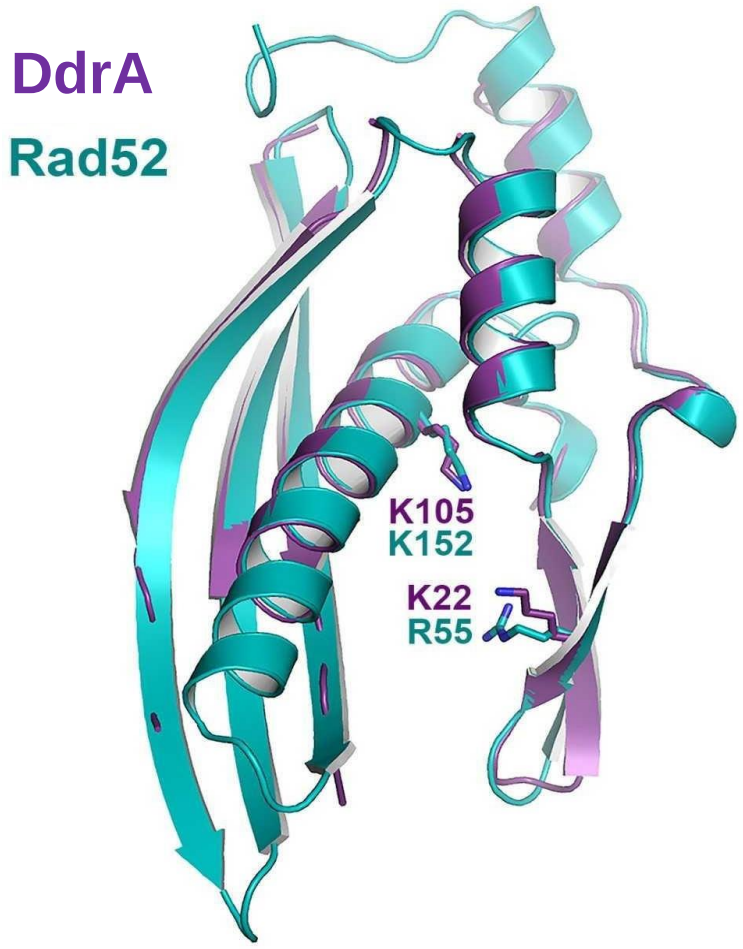


Figure 2.4.2: Mutant Design: Thread-Based Homology Modelling. Residues K22 and K105 of DdrA were found to be respectively analogous to R55 and K152, two residues that Rad52 requires for DNA annealing. The two basic residues were mutated to non-polar alanine residues in an effort to inhibit the postulated interaction with the negatively charged backbone of DNA.

As illustrated by Figure 2.4.3, full-length DdrA$^{K22A/K105A}$ was incapable of binding ssDNA. There was no observable ssDNA binding even at 20 μM, which was 10-fold higher than the concentration required for wild-type DdrA to fully bind DNA. DNA binding was not observed for poly-thymine or poly-adenine substrates of any length tested. Since ssDNA binding is required for annealing, it was therefore anticipated that the mutant would fail to enhance SSA. As expected, DdrA$^{K22A/K105A}$ failed to promote annealing above control levels (Figures 2.4.4 & 2.4.5). These findings suggest that DdrA requires K22 and K105 to first bind and subsequently anneal ssDNA.

TEV-digested DdrA$^{K22A/K105A}$ was found to adopt a heptameric arrangement by SEC analysis, suggesting proper folding at the monomeric level, and in turn, at the quaternary level. This finding is significant as it suggests that the lack of activity was a direct result of the chemical disruption of the active site as opposed to misfolding, which may have arisen from the substitution of the two hydrophobic alanine residues in place of the two positively charged lysine residues.

To test the biological significance of DdrA annealing, a DdrA knockout was complemented with wild-type or mutant DdrA and survival monitored in response to increasing concentrations of mitomycin C. Whereas 80% of wild-type cells survived exposure to 100 ng/mL of MMC, DdrA knockout and DdrA$^{K22A/K105A}$ complemented knockout cells failed to survive under these conditions (Figures 2.4.6 & 2.4.7), suggesting that ssDNA annealing by DdrA is essential for DNA damage tolerance *in vivo*.

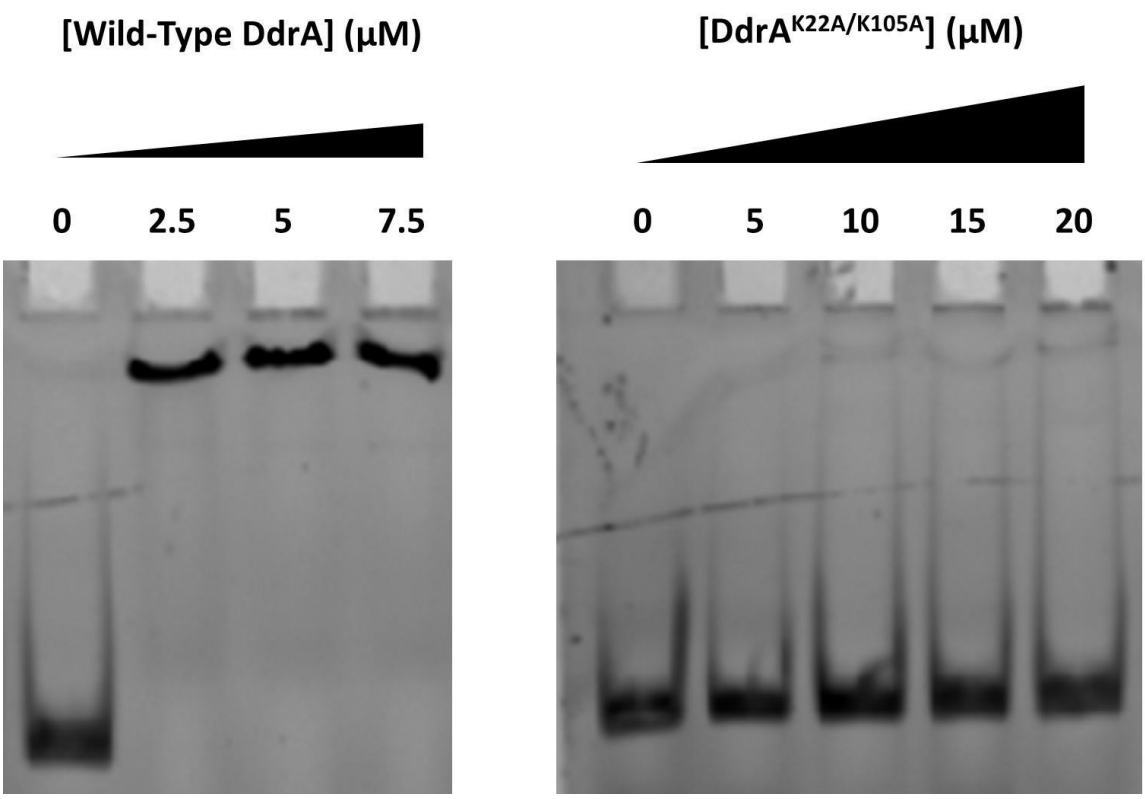


Figure 2.4.3: ssDNA Binding Assessment of DdrA$^{K22A/K105A}$. Successful ssDNA binding by full-length, wild-type DdrA (left) is shown in comparison to failed ssDNA binding by DdrA$^{K22A/K105A}$ (right). The results show that irrespective of the protein concentration, DdrA$^{K22A/K105A}$ is incapable of binding the DNA.

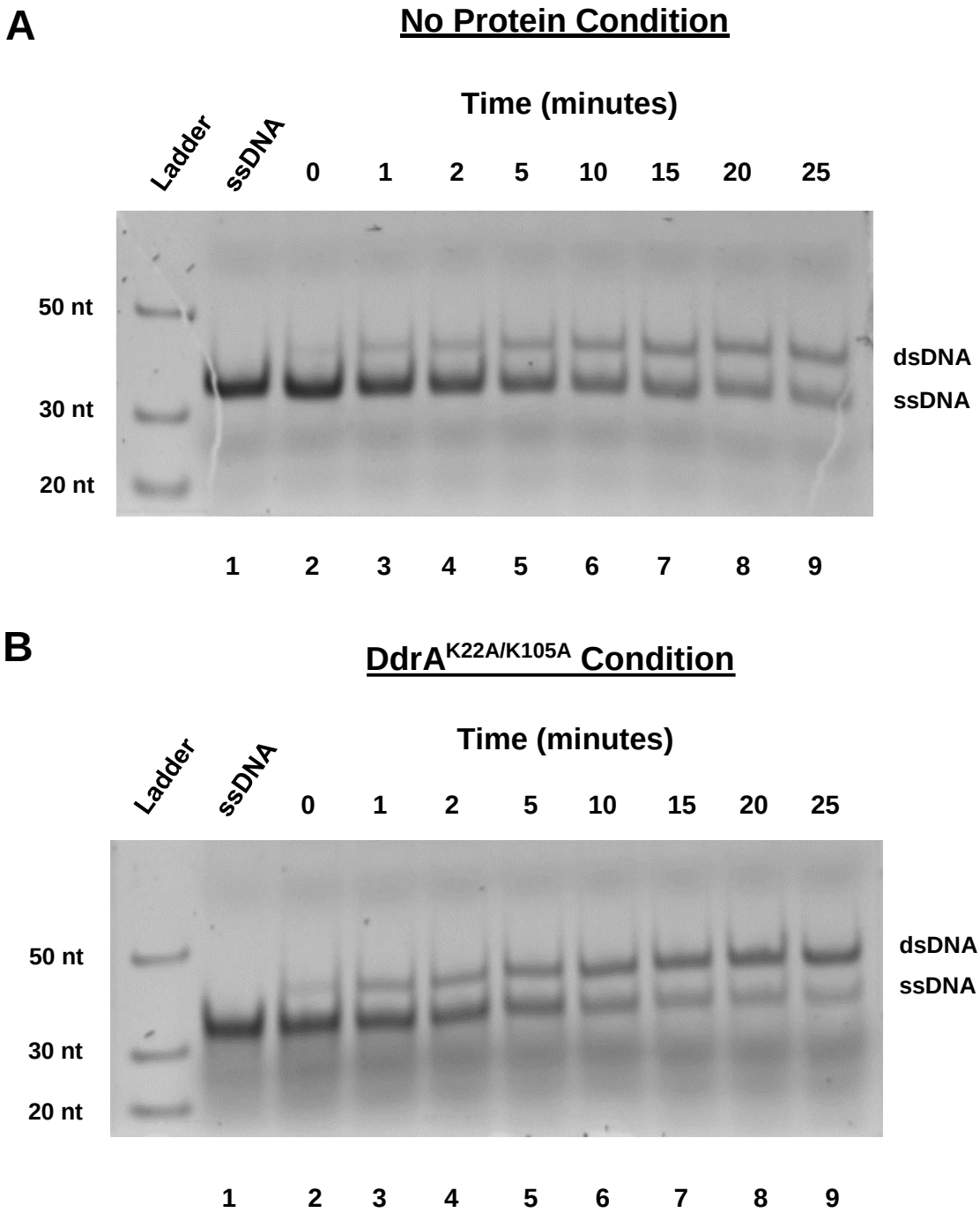


Figure 2.4.4: ssDNA Annealing Assessment of DdrA$^{K22A/K105A}$. ssDNA annealing is depicted both in the absence of protein (A) and in the presence of DdrA$^{K22A/K105A}$ (B). The results indicate that the mutated protein is incapable of enhancing the process of ssDNA annealing.

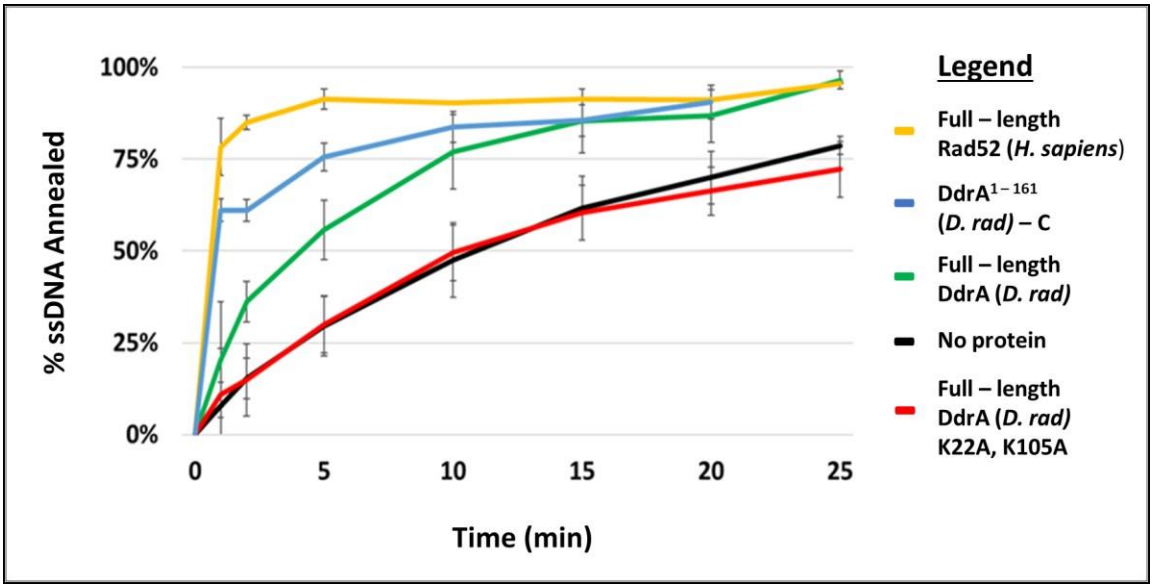


Figure 2.4.5: Comparison of Wild-Type and Mutant ssDNA Annealing. The ssDNA annealing results obtained for the mutated protein are compared to the results previously shown in Figure 2.3.4. The red line, representative of the mutant condition, is practically in-line with the black line, representative of the unassisted condition, suggesting that DdrA$^{K22A/K105A}$ is incapable of enhancing the process of ssDNA annealing.

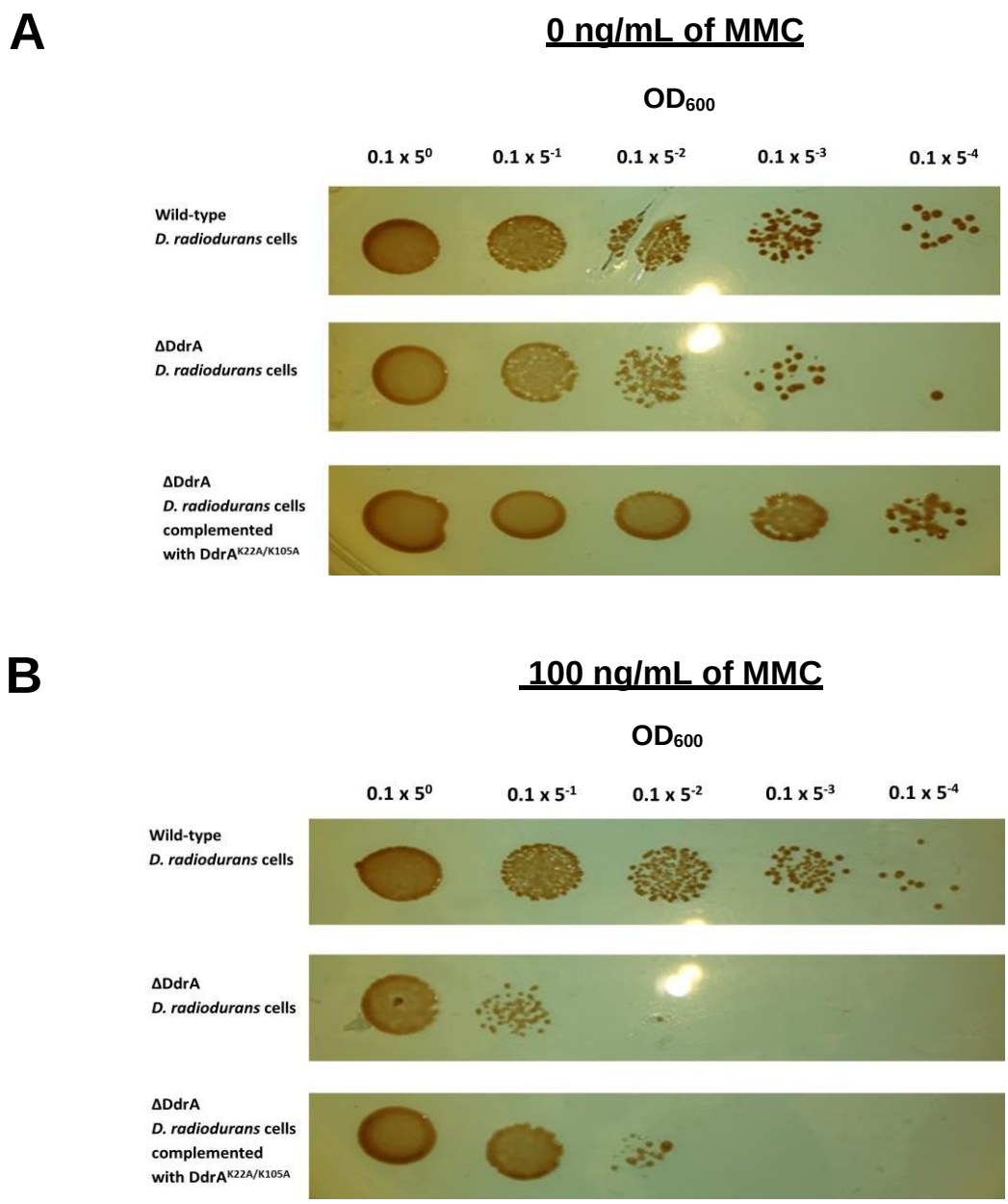


Figure 2.4.6: DNA Damage Repair via ssDNA Annealing: Pictorial View. The number of wild-type cells observed under normal conditions (A) appears unchanged following the introduction of 100 ng/mL of MMC (B), underscoring the resistance of *Deinococcus* to MMC. In contrast, the knockout cells died off almost entirely when exposed to this concentration of the drug, highlighting the importance of DdrA with respect to MMC-resistance. Complementing the knockout cells with DdrA$^{K22A/K105A}$ failed to restore resistance, suggesting that the annealing activity of the wild-type protein is important for extreme DNA damage tolerance *in vivo.*

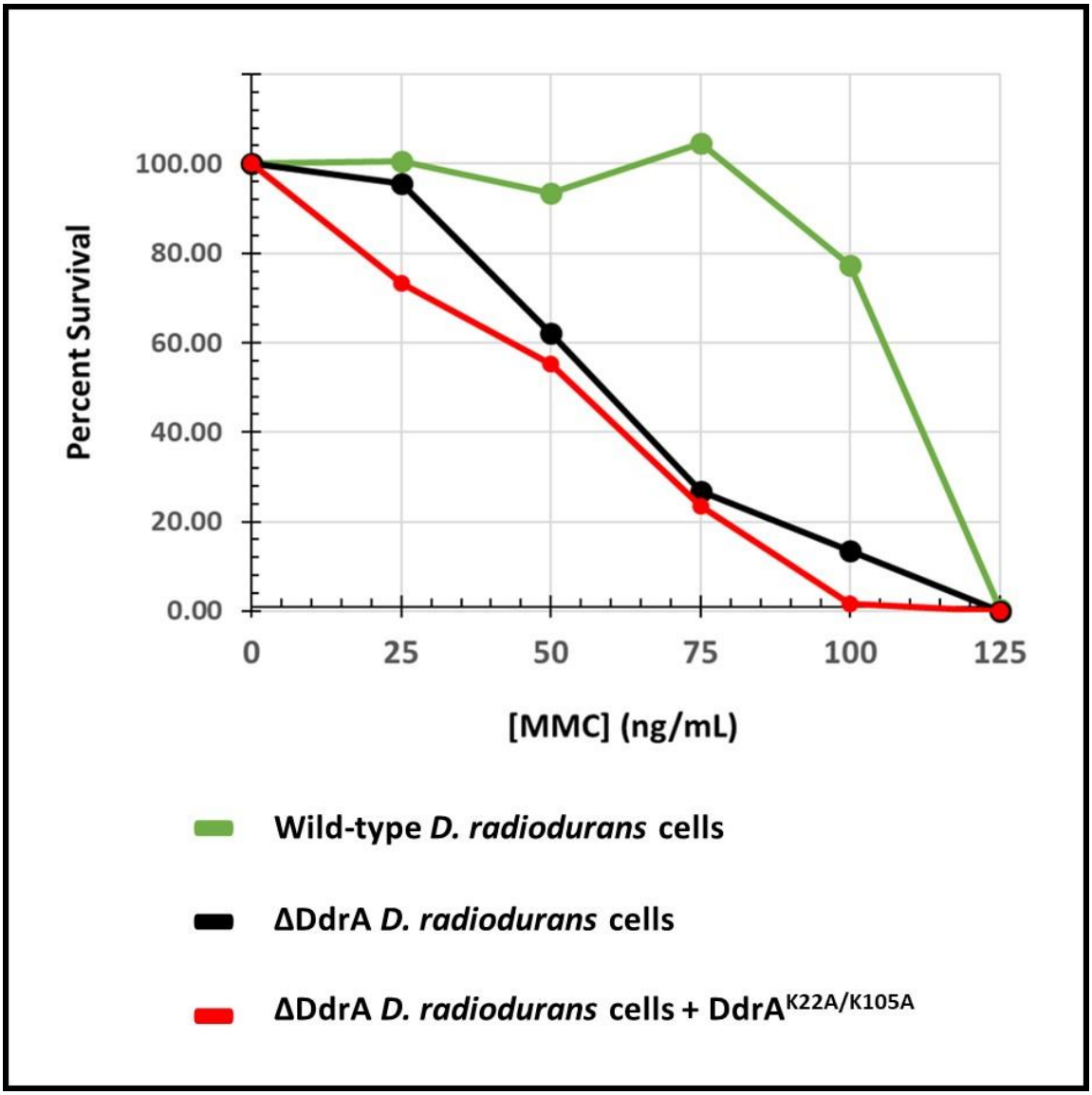


Figure 2.4.7: DNA Damage Repair via ssDNA Annealing: Graphical View. Wild-type *D. radiodurans* cells exhibited an extraordinary resistance toward mitomycin C, with no observable loss of viability up to a concentration of 75 ng/mL. Conversely, ΔDdrA *D. radiodurans* cells died off almost entirely when exposed to this concentration of the drug, confirming that DdrA is required for MMC-resistance. Complementing knockout cells with $DdrA^{K22A/K105A}$ failed to restore MMC-resistance, suggesting that the ssDNA annealing activity of DdrA is required for extreme DNA damage tolerance *in vivo*.

2.5 Discussion

Deinococcus radiodurans first rose to prominence when the bacterium was demonstrated to be capable of surviving extensive damage by a wide range of DNA damaging stimuli. Among the factors that have been identified to contribute to the DNA damage resistance profile of *Deinococcus* is the protein DdrA. Following extensive DNA damage, DdrA is upregulated at levels that surpass RecA. Furthermore, cells lacking DdrA are significantly more DNA damage sensitive than wild-type cells. Despite its essential role in DNA damage tolerance, prior to completion of work in this thesis, very little was known about the function of DdrA. With weak sequence similarity to hRad52 and characterized ssDNA binding activity, it seemed plausible that DdrA might function through ssDNA annealing. Our findings have validated this hypothesis and established a novel ssDNA annealing activity for DdrA. This activity was localized to an N-terminal domain, comprised of residues 1-161, and it appears to be partially regulated by elements in the less structured C-terminal region, comprised of residues 161-208. Two residues, K22 and K105, necessary for ssDNA annealing, were identified and used to establish a requirement for DdrA annealing activity in *Deinococcus* following exposure to extreme amounts of DNA damage. Altogether, this work not only suggests that DdrA functions as a Rad52 homologue for ssDNA annealing, but also represents the first identification of a functional prokaryotic Rad52 homologue.

2.5.1 DdrA: A Prokaryotic Rad52 Homologue

The fact that hRad52-like proteins are not widely conserved amongst prokaryotes had led to the suggestion that DdrA may serve a different function. This idea was supported by the finding that RecO is the functional homologue of hRad52 in bacteria. The notion was further strengthened by a report by Harris *et al*., which suggested that DdrA lacks ssDNA annealing activity (Harris *et al*., 2004). At the time, this finding was consistent with the idea that hRad52-like proteins likely serve alternative functions in prokaryotes (Iyer *et al*., 2002).

Work completed here indicates otherwise. DdrA does function as a ssDNA annealing repair factor, suggesting it should be included in the Rad52 superfamily of single-strand annealing proteins (SSAP's) despite very low sequence similarity (<10% identity). Initial sequence alignment of DdrA proteins from seven different strains of *Deinococcus* with Rad52 proteins from *T. aquaticus* and *H. sapiens*, coupled with thread-based homology modelling, identified two conserved basic residues, which we demonstrated have conserved function for ssDNA binding and annealing. Structural analyses have indicated that all members of the Rad52 superfamily of SSAP's adopt a similar fold. The thread-based homology model of DdrA (Figure 2.4.2) is indeed consistent with this fold. Taken together, these findings allow us to classify DdrA into the Rad52 superfamily of SSAP's. This classification represents a new outlook for the role of DdrA as it is in direct contradiction with a previous report, which claimed that DdrA lacks ssDNA annealing activity (Harris *et al*., 2004).

The Rad52 SSAP superfamily exhibits a sporadic phyletic distribution as members are notably absent from plants, nematodes and insects. Prior to our classification of DdrA as a member, the SSAP superfamily featured no prokaryotic members. Two additional superfamilies (RecT/Redβ and ERF) of ssDNA annealing proteins have been identified in prokaryotes. All three superfamilies are evolutionarily distinct with differing sequence conservation patterns and predicted folds. All members were originally bacteriophage proteins, which have since been adapted by a wide variety of evolutionarily distant cellular genomes (Iyer *et al*., 2002) for the common purpose of annealing ssDNA.

In *Deinococcus*, RecO has been shown to fulfill the role of HR initiation (see Section 1.4.2) and ssDNA annealing during homology-dependent and SSA repair. As such, RecO serves as the functional homologue of Rad52 in prokaryotes (Kantake *et al*., 2002) despite having a completely distinct tertiary structure from Rad52 (Makharashvili *et al*., 2004). This gave strength to the idea that another ssDNA annealing protein would not be required in *Deinococcus*. Our work here not only suggests that DdrA functions as a Rad52 homologue for ssDNA annealing, but also represents the first identification of both a functional and structural Rad52 homologue in prokaryotes.

2.5.2 The Role of DdrA in *Deinococcal* DNA Repair

As outlined in Section 1.5.1, DdrA is upregulated >20 fold following exposure to 3,000 Gy of ionizing radiation. This upregulation makes DdrA available to the cell for the purpose of DNA repair via ssDNA annealing. However, while DdrA

is required for the repair of extreme amounts of DNA damage, it is not required for normal cellular function. In fact, DdrA knockout cells are perfectly viable in the absence of extreme levels of DNA damage. This observation suggests that DdrA serves a highly specialized purpose. While the work presented here suggests that the specialized function in question is ssDNA annealing, it is unlikely that this is the only purpose of DdrA. This is supported by the finding that that the C- terminal region of DdrA is essential for DNA damage tolerance *in vivo*, but dispensable for ssDNA annealing *in vitro*. It will be important to further characterize the role of the C-terminus of DdrA in mediating protein-protein interactions between DdrA heptamers and perhaps other DNA repair factors.

While ssDNA annealing is unlikely to be the sole function of DdrA, this discovery does aid in understanding the overall placement of the protein in the *Deinococcal* repair pathway. As discussed in Section 1.4.5, knocking out DdrB in conjunction with DdrA results in greater radiosensitivity than knocking out either protein in isolation, suggesting redundant functions. Since work in the Junop lab has shown both proteins to be capable of ssDNA annealing (Sugiman-Marangos *et al*., 2016), it is therefore reasonable to suggest that DdrA and DdrB work together to enhance ssDNA annealing. Furthermore, knocking out DdrA or DdrB together with RecA, results in greater radiosensitivity than knocking out RecA alone. These findings suggest that the annealing activities of DdrA and DdrB occur independently of RecA. Indeed, ~ 30% of DNA breaks generated in response to extreme levels of damage are repaired in a RecA-independent manner. Taken together, these findings suggest that DdrA may function as a specialized ‘relief’

repair factor, able to repair DSB's through SSA annealing when RecA-mediated repair is overwhelmed.

Moving forward, it will be necessary to determine a high-resolution structure of DdrA, preferably in complex with DNA, to gain insight into the general mechanism of annealing. Since DdrA belongs to the Rad52 superfamily of SSAP's, the mechanism of ssDNA annealing by DdrA is likely comparable to that of Rad52. As available structures of Rad52 do not adequately inform on its annealing mechanism, a structure of DdrA bound to DNA could offer insight into the annealing mechanism of the entire Rad52 superfamily of SSAP's.

Recent work in the Junop lab has characterized the ssDNA annealing mechanism of DdrB (Sugiman-Marangos *et al*., 2016). Two pentamers bring together complementary strands of 30 nt DNA. Six bases are bound by each monomer, whereby four are exposed to the solvent and the other two are buried. Annealing is then enhanced via a two-step process: the exposed bases are first assessed for complementarity and if a match is found, the buried bases are then inverted and assessed in the same manner. Annealing is favoured in instances where perfect matches are detected. Both Rad52 and DdrA are believed to bind ssDNA in a similar fashion: 4 bases are bound per monomeric copy of heptamer (Parsons *et al*., 2000). Given these commonalities, it is therefore plausible to suggest that all three proteins may anneal ssDNA via comparable mechanisms. If future research determines this to indeed be the case, a universal mechanism by which ssDNA annealing proteins function will have been uncovered., which will represent a major step forward in understanding DNA repair in many organisms.

Chapter 3: Structural Characterization of DdrA

3.1 Abstract

The genome of *Deinococcus radiodurans* is quickly and faithfully restored following fragmentation induced by DNA damaging stimuli. Work presented in Chapter 2 demonstrated that DdrA contributes to the restoration of the *Deinococcal* genome by enhancing strand annealing. To gain more insight into the mechanism of strand annealing, we sought to determine an X-ray crystal structure of a DNA-bound DdrA complex. This chapter summarizes the progress made in this regard, in particular, with respect to identification of DdrA domain boundaries that are required for successful expression, purification and crystallization of the protein. Although the work has led to identification of conditions that generate high-quality diffracting crystals, a structure of DdrA was unable to be determined and will require continued work to optimize crystals for selenomethionine derivatization that can be used for phasing. As such, the chapter ends with recommendations for future efforts that may be helpful for the elucidation of a high-resolution structure.

3.2 Introduction

DdrA has been implicated in the extraordinary resistance of *Deinococci* to DNA damaging stimuli. This is best demonstrated by the findings that DdrA is upregulated >20 fold following exposure to 3,000 Gy of gamma radiation and that cells lacking DdrA are highly radiosensitive. As such, DdrA has been the focus of great interest within the field. Our recent finding that DdrA functions as a single-strand annealing factor opens new questions about its precise role in damage tolerance and its overall mechanism of action. Currently, there are no high-

resolution structural data for DdrA. DdrA$^{1-160}$ (*D. deserti*) has been characterized at low-resolution using negative stain EM. The 23 Å reconstruction revealed a surprisingly complex quaternary structure. DdrA was found to assemble into a heptameric ring that further trimerized to form an arrangement comprised of 21 individual subunits. The functional and biological relevance of this large complex have yet to be determined. Obtaining a high-resolution structure of the protein through X-ray crystallography, especially in complex with DNA, would offer valuable mechanistic insight into structure-function relationships and provide the basis for new mechanistic studies. This chapter outlines the significant progress made toward achieving this goal. In particular, DdrA domain boundaries, necessary for successful expression, purification and crystallization of the protein, are reported. Suggestions are provided for further efforts to complete the structure of DdrA.

3.3 Materials & Methods

3.3.1 Secondary Structure Predictions & Homology Modelling

Position-**S**pecific **I**terative-Basic Local Alignment Search Tool Based Secondary Structure **PRED**iction (PSIPRED) was used to predict the secondary structures of DdrA homologs from nearly thirty different strains of *Deinococcus* in an effort to determine the most suitable candidates for crystallization (see Figure 3.1.1). Phyre2 was also used to model tertiary structures of DdrA monomers. These structures contained most of the core structural elements found in the annealing domain of hRad52. There were, however, notable differences in the

lengths of secondary structure elements and loop regions. Taken together, these analyses helped guide the choice of DdrA homologues and domain boundaries used for structural studies (see Figure 3.1.2).

3.3.2 Protein Preparation

Genes encoding different DdrA homologues and truncations were codon optimized (GenScript) for expression in *E. coli* and gateway cloned from pUC57 entry vectors into destination expression vectors encoding either an N-terminal His_6 tag and TEV protease cleavage site (pDEST527) or C-terminal uncleavable His_6 tag (pDEST14). The integrity of final expression vectors was verified by Sanger sequencing.

Proteins were expressed in *E. coli* BL21 (DE3)-T1R cells (Invitrogen™) grown in LB supplemented with ampicillin (100 µg/mL) at 37°C to an OD_{600} of approximately 0.5 and induced with 1 mM of IPTG for 16 hours at 16°C. Cells were harvested by centrifugation at 6000 x g for 15 min at 4°C. Cell pellets were resuspended in lysis buffer (1 M NaCl, 20 mM Tris-HCl pH 8.0 and 5 mM imidazole) using 10 mL buffer per gram of cell pellet and lysed by 4 sequential French press passages at 10,000 psi. Following clarification by centrifugation at 48,384 x g for 40 min, soluble lysate was loaded onto a Ni-charged HisTrap™ FF 5 mL column (GE Healthcare) at 1 mL/min using an ÄKTA FPLC system. The column was washed with 15 column volumes each of buffer (1 M NaCl and 20 mM Tris-HCl pH 8.0) containing increasing amounts of imidazole (0, 5, 15, 30, 45 and 60 mM imidazole) prior to elution (600 mM imidazole). Eluted protein was buffer

exchanged using a HiPrep 16/10 desalting column (GE Healthcare), equilibrated with either TEV protease buffer (150 mM NaCl, 50 mM Tris-HCl pH 8.0, 5 mM DTT and 0.5 mM EDTA) or storage buffer (1 M NaCl, 20 mM Tris-HCl pH 8.0, 10% glycerol [v/v] and 5 mM imidazole).

While the C-terminal His_6 tag of pDEST14 constructs was designed to be uncleavable, the N-terminal His_6 tag of constructs cloned into pDEST527 was able to be removed by digestion with TEV protease. TEV cleavage reactions were performed with a 10:1 ratio of fusion protein to protease. Following digestion with TEV protease, samples were exchanged into lysis buffer (1 M NaCl, 20 mM Tris-HCl pH 8.0, 10% glycerol [v/v] and 5 mM imidazole) and passed over a 5 mL Ni-charged HisTrap™ column to isolate digested protein. Proteins were concentrated by ultrafiltration (10k MWCO, VivaSpin), aliquoted and stored at -80°C.

3.3.3 Analysis of Quaternary Structure

To assess monodispersity of purified DdrA and estimate quaternary structure, SEC was performed. DdrA (10 mg/mL) was resolved on a HiLoad™ 16/60 Superdex™ 200 prep grade column (GE Healthcare) using an ÄKTA Pure system (GE Healthcare) housed at 10°C. The column was equilibrated and run with buffer containing 20 mM Tris-HCl pH 8.0, 1 M NaCl and 15% glycerol (v/v). Molecular weight standards were run under the same conditions to calibrate the column for size estimation of DdrA.

SEC-MALS was also performed under the same conditions with the SEC column connected in-line to a Dawn HELEOS II MALS detector equipped with a 662 nm laser source and Optilab T-rEX differential refractometer equipped with a 658 nm LED source (Wyatt Technology, Santa Barbara, CA, USA). Molecular weights were calculated by Zimm plot analysis using ASTRA software (v6.1.5.22; Wyatt Technology).

To further assess quaternary structure, analytical ultracentrifugation (AUC) was used to perform a sedimentation velocity experiment. Prior to AUC analysis, an absorbance spectrum for DdrA was generated (Nanodrop Microvolume Spectrophotometer) in an effort to determine the most optimal parameters for AUC. AUC was performed at 20°C using a DdrA concentration of 0.25 mg/mL and a gravitational force of 11,612 x g (tracked at 250 nm). DdrA was suspended in buffer containing 150 mM NaCl, 30 mM Tris-HCl pH 8.0 and 10% glycerol (v/v); the dynamic viscosity (η) and density (ρ) of which were computed using Sednterp software. The partial specific volume ($\vec{v}$) of the protein was computed (Sednterp) based on its amino acid composition.

3.3.4 Crystallization

ssDNA Synthesis

Oligonucleotides used in crystallization trials were purchased unlabelled and PAGE purified from BioBasic. With the exception of a pair of mismatched oligonucleotides (5'-TGCT**T**GCT**T**GCT**T**GCT**T**GCT**T**GCT**T**GCT-3'; 5'-AGC**T**AGC **T**AGC**T**AGC**T**AGC**T**AGC**T**AGCA-3'), all oligonucleotides were poly A or poly T.

Broad Screening

Preliminary crystallization trials were conducted with DdrA in the absence of DNA. Using the hanging drop vapor diffusion method, 1 μL of protein was mixed with 1 μL of a mother liquor and dehydrated over 1 mL of 1.5 M ammonium sulfate. Mother liquor was obtained from commercially available kits, each with 96 individual conditions (Wizard I & II, Pact Premier HT-96, Hampton 110 & 112, MSG III, Nextal-AmSO_4, MCSG II, Nextal-PEGS, Sigma Basic, Sigma Low Ionic Strength, PEGRx-HT, MPD Suite, Morpheus I, Helix^{TM}-96, Kerafast, JBS, MD Midas, BioGenova and MSG IV). Crystallographic trays were incubated at either 20°C or 4°C and periodically examined for crystal growth. Table 3.2 shows a summary of all crystallographic trials.

Optimization

Crystals obtained from initial broad screening were confirmed to be protein crystals by collecting X-ray exposures (MicroMax-007 HF X-ray generator housed with a Saturn 994+ high resolution CCD detector). Promising conditions that generated confirmed crystals were optimized in an effort to obtain crystals with more favorable diffraction resolution. Optimization trays were designed by varying concentrations of each component in the original mother liquor along with protein concentration. Secondary optimization involved varying the concentration of ammonium sulfate (i.e. 0.75, 1, 1.25, 1.5,1.75 and 2 M) as a function of drop ratio (i.e. 1:2, 1:1 and 2:1).

Co-crystallization

Once the protein-only conditions were adequately assessed, DNA was added in a 1.2:1 molar ratio relative to the protein and crystallographic trays were set, as before. Medium sized oligonucleotides were prioritized with respect to the trials attempted since EMSA analysis had determined that these interact favorably with DdrA. ssDNA oligo's tested varied in size and included the following lengths: 14, 21, 28, 29, 30, 35, 42, 49 nt.

X-ray Diffraction Data Collection

Promising crystals were mounted on cryoloops™ (Hampton Research) and flash frozen directly in the nitrogen stream of a cryojet (Oxford Cryosystems), maintained at a temperature of 100 K. Initial screening was performed using a MicroMax-007 HF X-ray generator (Rigaku), equipped with VariMax optics and a Saturn 994+ high resolution CCD detector. Higher resolution data sets were collected at the Advanced Photon Source (APS) in Argonne, Illinois, USA as well as at the Canadian Light Source (CLS) in Saskatoon, Saskatchewan, Canada.

3.4 Results

3.4.1 DdrA Construct Design

Since crystallization is highly dependent on the ability to stably pack protein units in 3-dimensional arrays, it is essential to use protein that is well folded. Most proteins are arranged in domains with varying amounts of less structured intervening regions. By carefully selecting domain boundaries, it is often possible

to achieve higher degrees of success in crystallization. As such, considerable attention was paid in selecting domain boundaries for DdrA structural studies.

A combination of multiple sequence alignment, secondary structure prediction and homology modelling was used to determine initial domain boundaries. Sequences of DdrA from 28 different species of *Deinococcus* were subjected to secondary structure analysis using PSIPRED. Although most species were found to have highly conserved secondary structural elements, some variation occurred in the C-terminal regions, suggesting slightly different folding. To ensure maximal coverage, two homologues (*D. radiodurans* & *D. swuensis*) were chosen that represented these differences. DdrA from *D. radiodurans* was selected since it represented homologues with the least structured C-terminal domain and is the most highly studied homologue within the literature. Conversely, DdrA from *D. swuensis* was chosen since it had the most ordered C-terminal region.

As shown in Figure 3.1.1, DdrA was predicted to be largely structured with the exception of an extended unstructured region between residues 118 and 151. To further assess the importance of this region for structural integrity, homology modelling was performed using Phyre2. Homology models of DdrA were then compared to the structure of hRad52 (Figure 3.1.2). Interestingly, the region of DdrA predicted to be unstructured corresponded to an important segment of hRad52 needed to stabilize inter-subunit protein-protein interactions (labelled as the ’oligomerizing motif’ in Figure 3.1.2). Although the region was poorly modelled

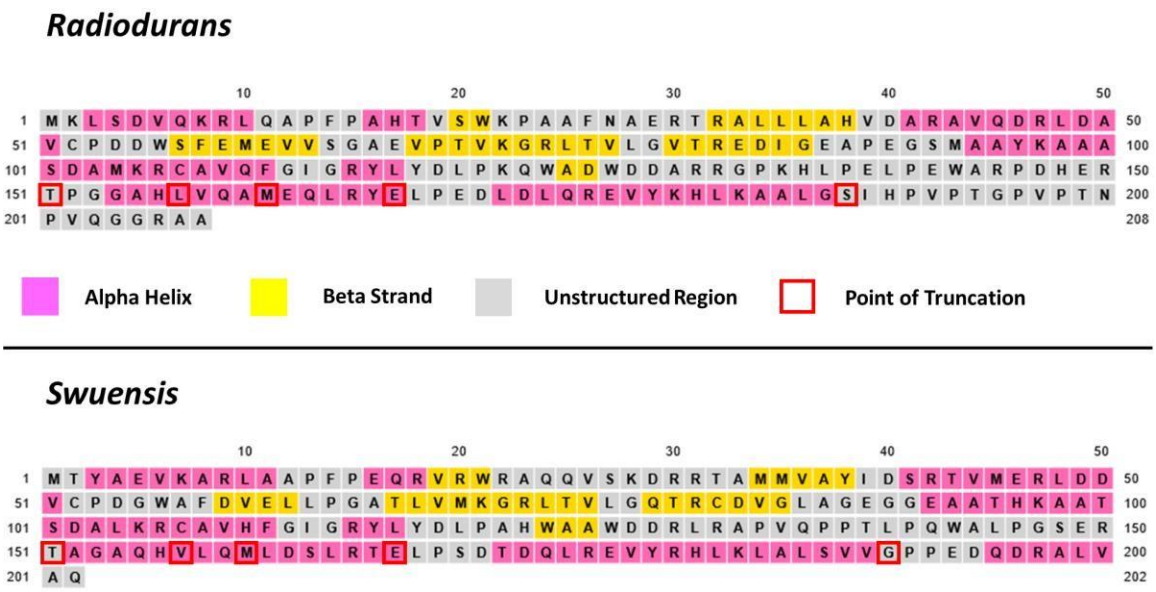


Figure 3.1.1: Secondary Structure Predictions for *Deinococcal* DdrA. The secondary structure predictions for DdrA proteins from *Deinococcus radiodurans* and *swuensis* are depicted. Residues predicted to exist in alpha helical conformations are highlighted in pink, beta strands in yellow and unstructured regions in grey. Residues that served as the boundaries for truncations are outlined in red.

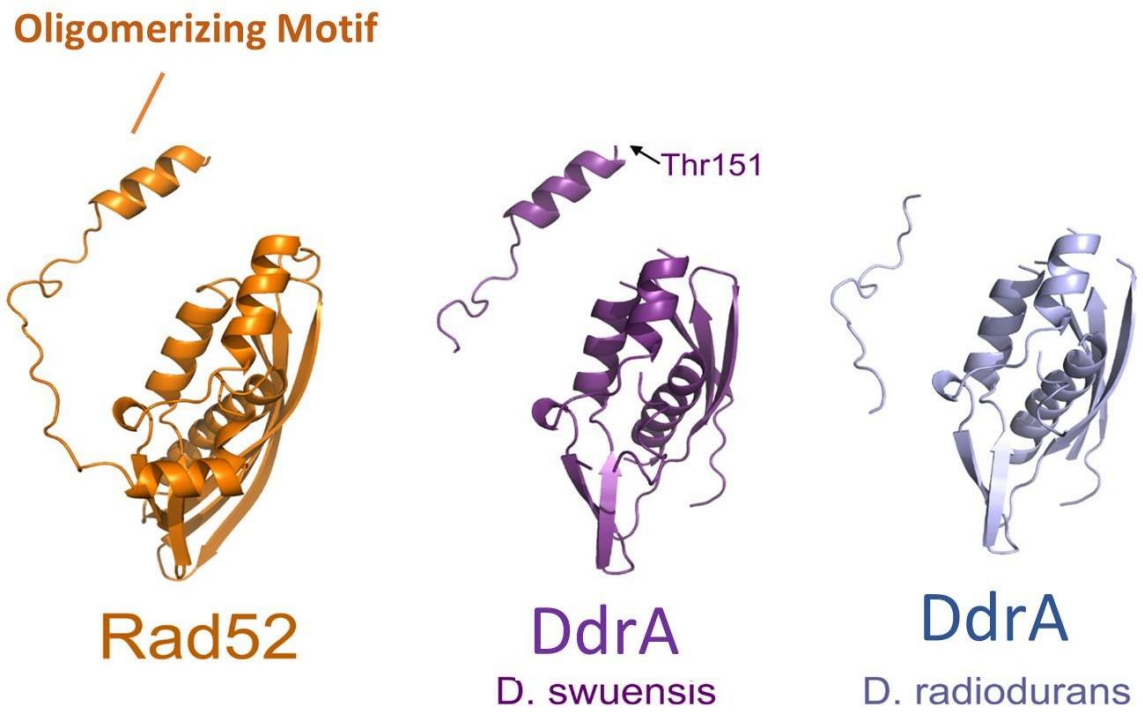


Figure 3.1.2: Homology Modelling of DdrA. Using Phyre2, it was found that the first 151 residues of DdrA from both *Deinococcus radiodurans* and *swuensis* could be modelled into a small domain with structural similarity to hRad52. The oligomerizing motif, which is believed to aid inter-subunit protein-protein interactions, is annotated.

within DdrA, we decided to keep it for all truncations to ensure proper oligomerization. Unfortunately, it was difficult to choose a precise domain boundary just beyond this region since the helix spanning residues 154 to 167 was not predicted with a high degree of confidence. Therefore, a number of different boundaries were initially designed spanning this region (depicted in Figure 3.1.1 by red squares). Since the N-terminal region of DdrA was predicted to be highly structured and part of the Rad52-like domain, all constructs retained the native N-terminal sequence. The first C-terminal boundary was chosen at residue 151, corresponding to the shortest sequence able to be modelled by Phyre2. Boundaries at 157, 161 and 167 were designed to cover an even span of the poorly predicted helix corresponding to the last secondary structure modelled by Phyre2. The final boundary at residue 188 was chosen because it retained the final well-conserved helix while removing the remainder of the C-terminus.

The secondary structure predictions by PSIPRED appear to be in agreement with the tertiary structure predictions by Phyre2. The alpha helical and beta sheet conformations, which were predicted to exist within the Rad52-like domain of DdrA by PSIPRED, feature prominently in the homology model generated by Phyre2.

3.4.2 Protein Preparation

Each DdrA construct was expressed and purified in parallel. Although all but one of the constructs expressed at high levels, several proteins could not be recovered in soluble form (Table 3.1). The only construct of DdrA from *D. swuensis* that expressed a significant amount (>1 mg of protein per L of cell culture) of soluble protein was the full-length. Since a favourable purification protocol had been developed previously (see Chapter 2) the protocol was not further optimized (i.e. altering cell lines for growth or destination vectors for expression). Instead, these truncations of DdrA from *D. swuensis* were deprioritized for crystallographic purposes. As mentioned, all truncations of DdrA from *D. radiodurans* exhibited high solubility, leading to excellent purification yields (~2-5 mg per L of cell culture).

***Deinococcal* DdrA**	
Radiodurans	***Swuensis***
1-151	1-151
1-157	1-157
1-161†	1-160
1-167	1-167
1-188	1-190
Full-length	Full-length

Legend

- Soluble: High Purification Yields
- Insoluble: Low Purification Yields
- Uninducible: No Expression Observed

† Refers to both N and C-terminally polyhistidine-tagged constructs

Table 3.1: Summary of Protein Properties. All *D. radiodurans* constructs exhibited high levels of solubility, leading to excellent purification yields. In constrast, all of the *D. swuensis* constructs, aside from the full-length protein, were either insoluble or unable to be expressed.

Although the His_6 tag was removed from DdrA for functional studies, we were unable to generate enough fully cleaved DdrA for structural studies. Despite being able to achieve ~75% cleavage efficiency when large amounts of TEV protease were added, isolation of fully cleaved heptamers proved challenging using IMAC, SEC or ion exchange chromatography. Therefore, structural studies were performed with constructs that retained uncleaved N-terminal His_6 tags.

3.4.3 Crystallization

DdrA proteins at varying concentrations (2-10 mg/mL) were initially used for broad screening to identify conditions able to induce crystal formation. Using the hanging drop vapor diffusion method, equal volumes of protein and mother liquor were mixed and dehydrated over 1.5 M ammonium sulfate at 4°C and 20°C. All crystallization trials are summarized in Table 3.2. Broad screening of protein only conditions resulted in many crystals, however, very few were confirmed to be protein following exposure to X-rays.

Construct	+/- His?	DNA Added	Number of Drops Set	Temp's	Results
Full-length DdrA (*D. rad*)	+ His	None	3 x 96 = 288	20°C	No crystals
Full-length DdrA (*D. swu*)	+ His	None	23 x 96 = 2, 208	20°C	No crystals
DdrA$^{1\text{-}151}$ (*D. rad*)	+ His	None	9 x 96 = 864	4°C, 20°C	No diffracting crystals
DdrA$^{1\text{-}160}$ (*D. rad*)	+ His	28 nt Poly T 29 nt poly T	13 x 96 = 1,248	20°C	Several diffracting crystals
DdrA$^{1\text{-}161}$ (*D. rad*) N-terminal His-tag	+ His	28, 29, 30, 35, 42, & 49 nt Poly A 14, 21, & 28 nt Poly T Mismatched DNA	19 x 96 = 1,824	4°C, 20°C	6 crystals with poor diffraction
DdrA$^{1\text{-}160}$ (*D. geothermalis*)	+ His	None	4 x 96 = 384	20°C	2.4 Å resolution
DdrA$^{1\text{-}167}$ (*D. rad*)	+ His	28 nt Poly T	2 x 96 = 192	20°C	No diffracting crystals
DdrA$^{1\text{-}188}$ (*D. rad*)	+ His	Mismatched DNA	3 x 96 = 288	20°C	No diffracting crystals
Total: 8 constructs	+ His	12 total oligonucleo-tides	76 x 96 = 7,296	4°C, 20°C	2.4 Å resolution

Table 3.2: List of Crystallographic Trials.

In parallel, we carried out crystallization trials of DdrA in the presence of ssDNA. Addition of ligands (such as DNA) frequently causes proteins to adopt more constrained states that are more amenable to crystallization. Since the exact number of bases bound by each DdrA subunit was unknown, a number of different sized oligonucleotides were used. These ranged in size from 14 to 49 nt, in multiples of 7; however, the majority of trials were performed with medium sized oligos (~28 nt) since EMSA analysis suggested saturated binding at this length. Co-crystallization with $DdrA^{1-161}$ resulted in identification of several conditions that produced protein crystals of similar morphology. These crystals appeared as hexagonal rods of varying dimensions (Figure 3.2 B). Each of the conditions that produced these crystals had $NH_4H_2PO_4$ in common. We therefore optimized $NH_4H_2PO_4$ concentration, drop ratio, drop size and dehydrant concentration to improve crystal size and quality. Although crystals of sufficient size (100 x 30 x 30 μm) could be obtained following optimization, diffraction quality remained poor. Even with long exposures, these crystals failed to diffract X-rays to more than 15 Å resolution. Crystals generated with protein-DNA complexes frequently suffer from this problem since DNA ends are not well-ordered. To overcome poor diffraction, ssDNA length was varied slightly (28, 29, 30 nt); however, crystal diffraction could not be further improved. With limited options to pursue, we considered the possibility that the presence of an N-terminal His_6 tag might be limiting crystal packing, resulting in poor diffraction.

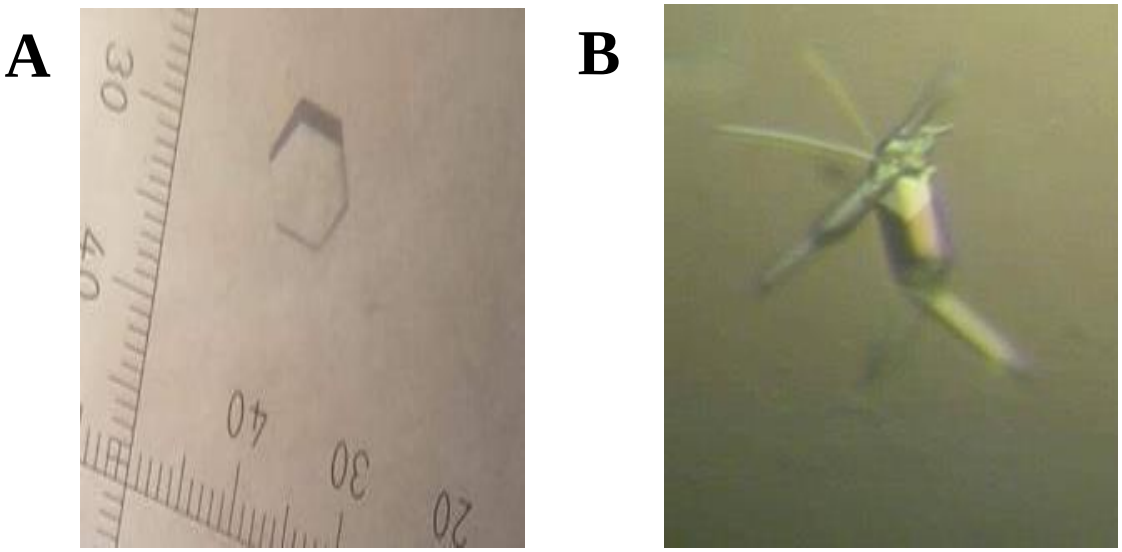


Figure 3.2: Examples of Crystals Obtained. (A) This crystal was obtained following the optimization of an earlier crystal, which was believed to be comprised of DdrA$^{1-151}$ (*D. rad*). Upon screening, the crystal above produced a diffraction pattern indicative of salt and as such, is an example of a false positive result. (B) This crystal was obtained following the co-crystallization of DdrA$^{1-161}$ (*D. rad*) with 28 nt poly T DNA. Upon screening, the crystal failed to diffract to a high resolution, suggesting the need for further optimization.

To characterize the effect of an N-terminal His_6 on protein behaviour, SEC-MALS and AUC were performed on DdrA$^{1-161}$. Surprisingly, as shown in Figures 3.3.1 & 3.3.2, leaving the N-terminal tag intact resulted in DdrA$^{1-161}$ forming large heterogeneous entities (in the MDa range; mean molecular weight of 3.3 MDa), similar to what had previously been observed for full-length DdrA (Figure 2.2). Since it was not possible to remove the N-terminal tag efficiently enough to produce sufficient amounts of protein for structural studies, we chose to design additional constructs with C-terminal His_6 tags. Prior characterization of C-terminal

tagged DdrA$^{1-161}$ (Figure 2.2) indicated the presence of monodisperse, heptameric protein, suggesting this would be an effective approach.

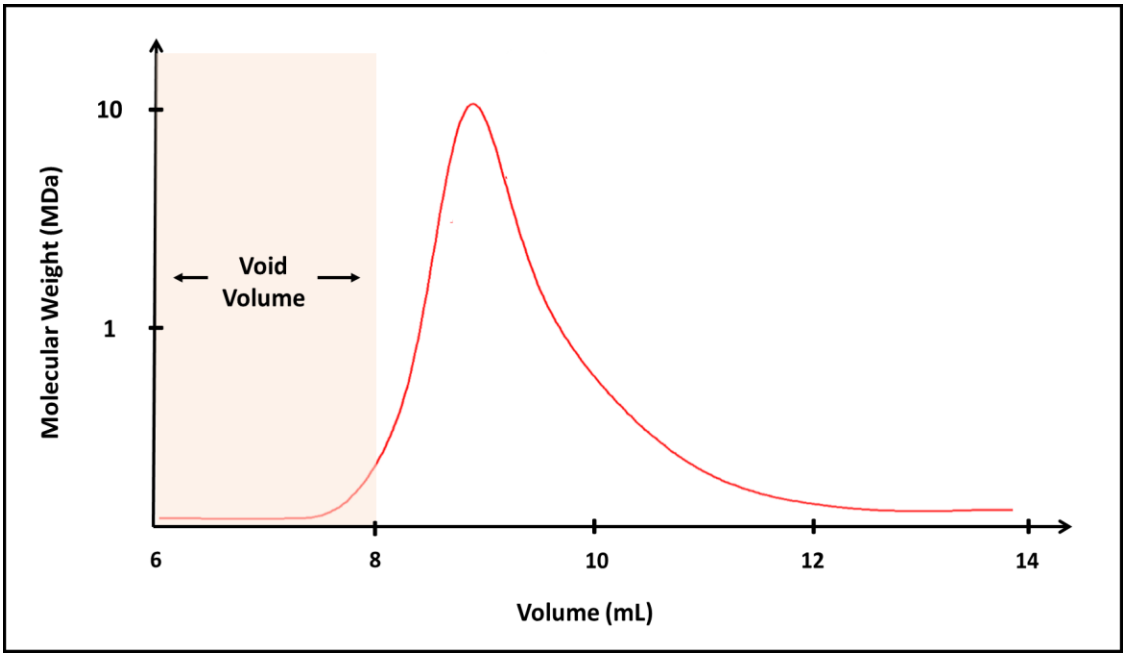


Figure 3.3.1: Analysis of Quaternary Structure of DdrA by SEC-MALS. A semi-logarithmic chromatogram obtained following analysis of the quaternary structure of N-terminally His_6-tagged DdrA$^{1-161}$ (*D. radiodurans*) by SEC-MALS. The analysis yielded inconclusive results, suggestive of a very large protein structure. Since the HiPrep^{TM} 16/60 Sephacryl^{TM} S-300 HR column is capable of resolving proteins up to 1.5 MDa in size, it would appear that the protein of interest is forming species of even higher molecular weight.

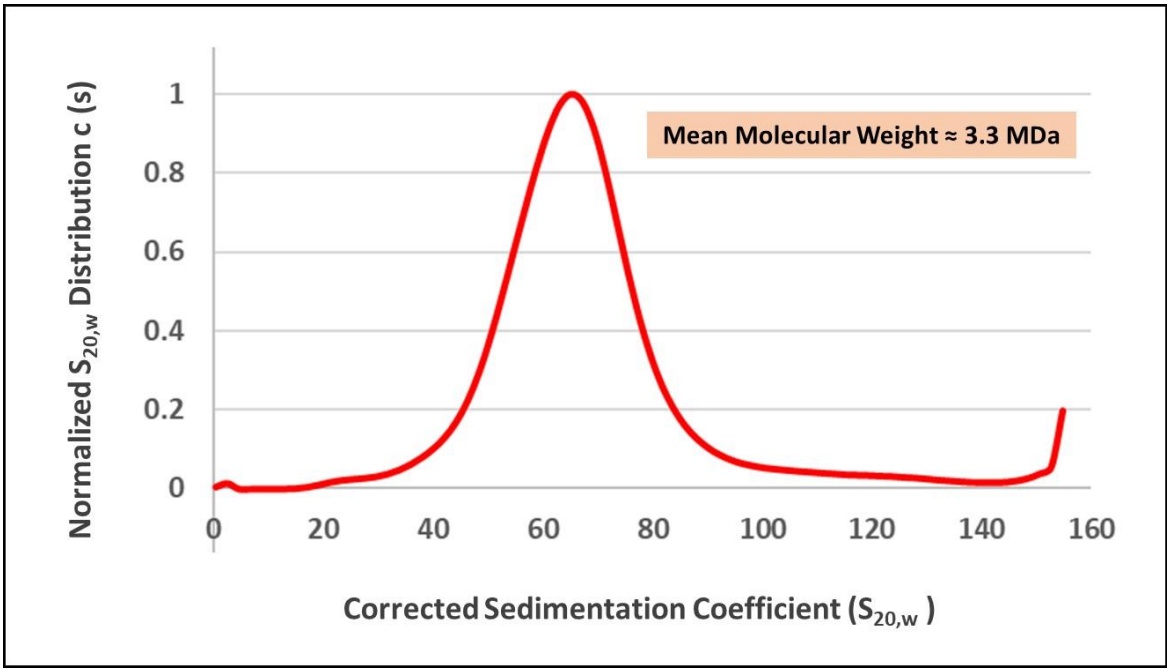


Figure 3.3.2: Analysis of Quaternary Structure of DdrA by AUC. Similar to analysis by SEC-MALS, AUC determined that N-terminally His_6-tagged $DdrA^{1-161}$ (*D. radiodurans*) forms very large complexes with a mean molecular weight of approximately 3.3 MDa.

Since $DdrA^{1-161}$ generated well-behaved protein with respect to crystallization, we therefore chose to explore the use of other DdrA homologues (*D. geothermalis* & *D. deserti*) with similar domain boundaries (residues 1-160 instead of 1-161). Using this approach, several protein crystals were obtained for $DdrA^{1-160}$ in the presence and absence of DNA (Figure 3.4); however, none diffracted to greater than 15 Å resolution.

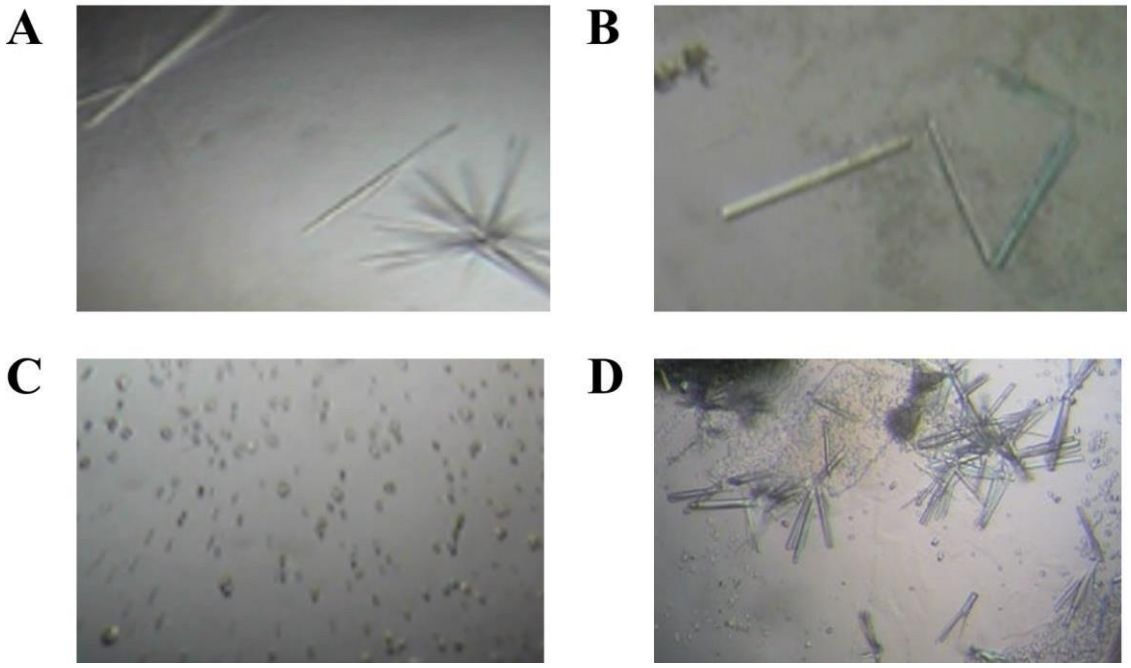


Figure 3.4: Crystals Obtained with DdrA$^{1-160}$ (*D. radiodurans*). Crystals depicted in (A) and (B) were obtained in the absence of DNA, whereas the crystals depicted in (C) and (D) were obtained in the presence of 28 nt Poly T and 29 nt Poly T, respectively. All crystallization trials were set at 20°C. Although all of these crystals were confirmed to contain protein, none diffracted to an appreciable resolution (i.e. >15 Å).

3.4.4 Preliminary X-Ray Diffraction of DdrA$^{1-160}$ (*D. geothermalis*)

During the writing of this thesis, the crystallographic component of the DdrA project was continued by two graduate students (Robert Szabla and Emily Pickering). One of the new DdrA$^{1-160}$ constructs from *D. geothermalis* resulted in the identification of several conditions able to generate protein crystals. One such condition yielded crystals (Figure 3.5) that diffracted to 2.4 Å resolution at the Canadian Light Source (CLS, Saskatoon, Saskatchewan). Although good quality

X-ray diffraction data were collected, the structure could not be determined by molecular replacement using hRad52 as a search model. Selenomethionine (SeMet) derivatized crystals have now been prepared and will be used to solve the structure via single-wavelength anomalous dispersion (SAD) phasing. The fact that hRad52 failed as a search model in molecular replacement suggests the similarity between DdrA and hRad52 may be less than originally anticipated.

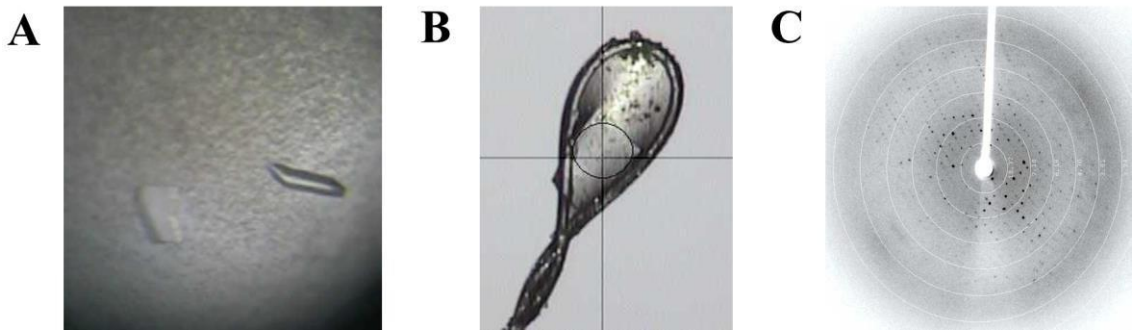


Figure 3.5: Preliminary X-Ray Diffraction of DdrA$^{1\text{-}160}$ (*D. geothermalis*). A crystal of DdrA$^{1\text{-}160}$ (*D. geothermalis*) is shown in the crystallization drop (A) and mounted on a 100-micron loop (B). The crystal initially diffracted to a resolution of 3.4 Å at the X-ray source located at Western University (C). However, a full data set was recently collected to 2.4 Å at the Canadian Light Source (Saskatoon, Saskatchewan).

3.5 Discussion

Functional studies performed in Chapter 2 revealed that DdrA is capable of ssDNA annealing enhancement. The successful identification of key residues for this activity led to the development of a non-functional mutant, which ultimately allowed us to demonstrate that the activity is important for DNA repair *in vivo*. Given the significance of these findings and given the fact that the molecular mechanism for any Rad52-related ssDNA annealing event had not yet been determined, further experimentation was deemed necessary. Our next step, therefore, was to probe the mechanism of ssDNA annealing at an atomic level by solving the structure of DNA-bound DdrA via X-ray crystallography.

Co-crystallization, as the name suggests, is a two-part problem. Both the correct oligonucleotide and the correct construct of DdrA must be carefully selected. Since our SEC analysis of $DdrA^{1-161}$ corroborated earlier EM studies suggesting a heptameric assembly of DdrA, we chose to use oligonucleotides lengths that were multiples of seven. We further chose Poly A and Poly T oligonucleotides, since these substrates are less likely to form secondary structures that might alter DdrA binding and/or orderly packing required for proper crystal formation. DdrA domain boundaries were chosen based on conservation of predicted secondary and tertiary structures. DdrA homologues from two *Deinococcus* species (*D. radiodurans* and *D. swuensis*), with slight variation in C-terminal structure, where further chosen for structural studies to improve crystallization probability.

Combined, these analyses uncovered important structural patterns conserved within DdrA proteins from 28 species of *Deinococcus.* For instance, the N-terminal region of DdrA (corresponding to residues ~1-151) was predicted to adopt a well-ordered Rad52-like domain (Figure 1.6). As discussed in Chapter 2, the annealing activity of this domain was found to be higher than that of the full-length protein and more in-line with Rad52 itself. Interestingly, a less structured, but highly conserved, region of DdrA (residues ~118-151) was identified that corresponded to the oligomerization motif of hRad52. Structural studies of hRad52 suggest that this motif is essential for assembly of a stable quaternary ring structure. In contrast, the C-terminal domain of DdrA was found to be less similar to hRad52 and less conserved amongst various DdrA homologues. Nevertheless, the C-terminal tails were predicted to be relatively well-ordered, suggesting structural conservation associated with mediating binding partner interactions (discussed in Chapter 2).

Our initial crystallization efforts, involving numerous constructs and conditions, led to very few crystals that could be positively identified as protein (see Figure 3.2 B). However, this crystal did not diffract well and efforts to optimize its quality failed. Our unexpectedly low success rate led us to reevaluate the domain boundaries and protein quality of constructs used for structural studies. Surprisingly, both SEC-MALS and AUC indicated that the N-terminally tagged version of $DdrA^{1-161}$ forms a large heterogeneous assembly, in the megadalton range. These results were similar to data obtained when full-length DdrA was subjected to both size-exclusion chromatography and SEC-MALS. Since these

properties are not ideal for crystallization, it not only explained our low success rate, but indicated that changes needed to be made.

Since previous experimentation indicated that C-terminal His-tagged DdrA could be produced as a homogeneous sample (~130 kDa), and that the N-terminal His-tag of DdrA constructs could not be efficiently removed by proteolytic cleavage, we decided to re-engineer DdrA constructs with uncleavable C-terminal His-tags. In addition to DdrA$^{1-161}$, several new constructs of DdrA were re-engineered to include the C-terminal tag, some with slightly altered domain boundaries. As well, additional homologues of DdrA (from *D. geothermalis* and *D. deserti*) were included to help improve crystallization success. These two strains were chosen based on their ability to grow most optimally around 50°C. Proteins from thermophilic organisms are typically more stable compared to those from mesophiles, often leading to improved crystallization success.

Preliminary crystallization attempts with C-terminal His-tagged DdrA$^{1-160}$ improved the number of crystallization hits (Figure 3.4), suggesting newly engineered DdrA constructs might be generally more amenable to crystallization. Indeed, recent studies (Robert Szabla and Emily Pickering in the Junop lab) using these constructs that were continued during the writing of this thesis, proved to be highly successful. Using re-engineered constructs, crystals (DdrA$^{1-160}$ [*D. geothermalis*]) were readily obtained that diffracted well using the home X-ray source. Although it is difficult to say which of the re-engineered changes (position

of His-tag, domain boundary, homologue) contributed most significantly to this success, this finding underscores the importance of their proper selection.

The crystal of DdrA$^{1-160}$ (*D. geothermalis*) diffracted to moderately high resolution (2.4 Å) using synchrotron radiation at the Canadian Light Source. Initial attempts to solve the structure of DdrA via molecular replacement using hRad52 as the search model were unsuccessful, suggesting structural differences between DdrA and hRad52. This is not unexpected given their low shared sequence identity (<10%). Moving forward, the plan is to use selenomethionine-derivatized crystals to solve the structure of DdrA using SAD phasing methods. Determining the structure of ssDNA-bound DdrA will provide significantly more mechanistic insight compared to the apo structure alone. Oligonucleotides have not yet been included with any of the thermophilic constructs for crystallization and so we remain optimistic that a ligand-bound structure will be obtained soon.

Chapter 4:

Summary &

Future Direction

4.1 DNA Damage Repair by *Deinococcus Radiodurans*

As the blueprint of life, DNA must be faithfully protected from damage. Reactive oxygen species routinely cause oxidation, which in turn leads to various DNA base modifications as well as single and double strand breaks. In addition to ROS, DSB's can arise in response to ionizing radiation, ultraviolet radiation, mitomycin C and desiccation. Unlike most organisms that only tolerate small amounts of DNA damage, *Deinococcus* is able to withstand extraordinary amounts. Interestingly, *Deinococci* accumulate DSB's to the same degree as radiosensitive organisms, indicating that damage tolerance does not result from prophylactic protection of the genome. Rather, the unique radioresistance is believed to stem from the action of several unique proteins. Like other proteins of this unique group, DdrA is strongly upregulated in response to extreme DNA damage and is essential for maintaining damage resistance. Although DdrA is thought to function in the context of ESDSA repair, its precise role has remained unclear.

Efforts to characterize a function for DdrA have provided limited insight thus far. Prior to work undertaken here, Harris *et al*. (2004) demonstrated that DdrA is capable of binding ssDNA, but not dsDNA; and that the N- and C-terminal regions are required for function *in vivo*. Although DdrA was reported to bear limited sequence similarity to Rad52, no ssDNA annealing activity was identified, suggesting an alternate role in repair. Nevertheless, since data were not provided to support a lack of annealing activity, we wished to confirm (or disprove) the

suggestion in order to better understand how DdrA might contribute to extreme DNA damage tolerance in *Deinococci*.

4.2 Summary of Findings

Our analysis of DdrA ssDNA annealing demonstrated that full-length DdrA is capable of strand annealing *in vitro*. Attempts to further localize this activity revealed that the C-terminal region of DdrA is dispensable for annealing. Surprisingly, in the absence of C-terminal residues (162-208), DdrA was found to have enhanced activity, suggesting a potential regulatory function within this region. Since annealing activity of $DdrA^{1-161}$ was comparable to hRad52, it seemed likely that DdrA annealing activity would be required for repair *in vivo*. To further explore this possibility, we identified conserved positively charged residues from a multiple sequence alignment of DdrA homolgoues and hRad52. In conjunction with structural comparison of hRad52 and DdrA (homology model), and subsequent *in vitro* analysis of a DdrA mutant, we were able to determine that K22 and K105 are essential for ssDNA annealing. Importantly, this allowed additional cell-based studies to be performed, which further suggested that ssDNA annealing by DdrA is required for DNA damage repair *in vivo.*

In an effort to elucidate the mechanism of ssDNA annealing, we sought to determine the structure of DdrA (alone and in complex with ssDNA) using X-ray crystallography. Preliminary trials, conducted in the absence of DNA, failed to generate many protein crystals. Subsequent trials included the addition of ssDNA. Oligo's differing in length by multiples of 7 nt were used in these trials based on the

finding that DdrA$^{1-160}$ (*D. deserti*) assembles into a heptameric ring (Gutsche *et al*., 2008). Co- crystallization trials resulted in several promising crystals; however, all failed to diffract to an appreciable resolution, which could not be further improved by optimization of growth conditions. With limited options, expression constructs were re-engineered incorporating changes to His-tag location, domain boundary and choice of species homologue.

Following my departure from the lab, continued crystallization studies using re-engineered DdrA constructs yielded many diffracting crystals. Therefore, the overall strategy used to generate DdrA crystals was proven effective. Crystals of DdrA$^{1-160}$ from *D. geothermalis* were recently used to collect a complete diffraction data set to 2.4 Å resolution (Canadian Light Source in Saskatoon, Saskatchewan). Although the data could not be used to solve the structure of DdrA by molecular replacement (using the hRad52 structure as a search model), selenomethione-derivatized crystals have now been prepared and will be used shortly to solve the structure using Se-SAD methods.

4.3 Implications of Findings

The finding that DdrA possesses ssDNA annealing activity allows us to classify the protein into the Rad52 superfamily of ssDNA annealing proteins. This classification is significant as DdrA is the first prokaryotic protein to become a functionally verified member.

In addition to DdrA, the *Deinococcal* protein RecO is capable of strand annealing; however, unlike DdrA, RecO adopts a completely different fold and exists as a monomer in solution (Makharashvili *et al*., 2004). The fact that RecO is capable of strand annealing despite not having an elaborate quaternary structure suggests that the primary mechanism used for strand annealing differs from DdrA. Although its mechanism remains unclear, the quaternary structure of DdrA may serve as a scaffold to allow C-terminal regions of DdrA to mediate interactions *in vivo*. This idea is supported not only by our work showing that the N-terminal 161 residues are required for annealing activity, but other studies demonstrating a requirement for the C-terminal region during repair *in vivo* (Harris *et al*., 2008). It is unlikely that radiosensitivity observed by Harris *et al*. (2004) was due to a lack of ssDNA annealing as we showed $DdrA^{1-161}$ to be perfectly capable of annealing. In fact, the truncated protein exhibited more robust annealing activity than the full-length protein, suggesting that the C-terminus may negatively regulate annealing. The finding that the C-terminal domain (dispensable for annealing) is required for DNA damage tolerance, suggests that DdrA is involved in other aspects of DNA repair, likely involving interaction with other repair factors.

While DdrA is likely responsible for multiple aspects of genomic restoration, the work presented here suggests a primary role in SSA. As outlined in Figure 1.4, SSA involves the formation of extended 3' ssDNA tails at the site of DSB's by the action of UvrD and RecJ. Although the majority of such intermediates are repaired by a RecA-dependent mechanism, approximately one-third are reassembled into larger fragments using single-strand annealing.

Prior studies in the Junop lab demonstrated that another repair protein, DdrB, is capable of enhancing SSA (Sugiman-Marangos *et al*., 2016), similar to DdrA. Interestingly, deletion studies involving RecA, DdrA and/or DdrB indicated that DdrA and DdrB have complementary activities in repair that are RecA-independent (Tanaka *et al*., 2004). This indicates that while both DdrA and DdrB function as ssDNA annealing proteins, they participate in separate repair pathways. Taken together, this further suggests that of the three ssDNA annealing factors identified in *Deinococcus* (RecO, DdrA and DdrB), each is involved in a separate repair pathway. Further studies will be required to identify specific partners (and mechanisms) involved in these pathways; however, at this point, it is clear that *Deinococcus* has evolved unique repair strategies for SSA repair that are essential for extreme DNA damage tolerance.

4.4 Future Direction

Obtaining the crystal structure of DdrA in complex with ssDNA would provide significant insight into the annealing mechanism. Recent successes in obtaining diffracting crystals of apo DdrA required iterative cycles of protein engineering. While this may be similarly required to obtain a structure of DNA-bound DdrA, it seems more likely that the limiting step will be identifying suitable DNA substrates. This should involve crystallization trials using a more focused length distribution of ssDNA, centred around 28 nt. Work here suggested that this is the approximate length for optimal interaction with DdrA and is in agreement with

prior studies involving hRad52 that demonstrated that each subunit of hRad52 binds 4 nt (Parsons *et al*., 2000).

How cells deal with the opposing need to prevent ssDNA from forming secondary structures and at the same time promote accurate strand annealing represents an important question in DNA repair. Overcoming the thermodynamic barrier of accurate annealing under biological temperatures requires annealing proteins that can provide 'proofreading' capability. This is particularly important for *Deinococcus* since response to extreme doses of DNA damage requires the cell to simultaneously orchestrate annealing for hundreds of fragments. Further studies are needed to understand how DdrA and Rad52 promote accurate ssDNA annealing. In combination with structural studies of ssDNA-bound DdrA, annealing assays could be performed to further probe mechanistic details. In particular, it will be important to establish the minimal requirements for accurate annealing by performing *in vitro* studies with carefully designed substrates, such as those containing differing numbers of mismatches at defined positions. Annealing assays, incorporating mismatched substrates could also be carried out with structure-guided mutational studies of DdrA. Together, such studies will help unravel the molecular mechanism for accurate ssDNA annealing.

Since DdrA has been shown to bear functional similarity to Rad52, it would be useful to examine whether DdrA is capable of promoting other types of ssDNA transactions (in addition to annealing) that Rad52 has been shown to be capable of executing. Rad52 is able to promote DNA strand exchange (Bi *et al*., 2004),

strand invasion (Lok *et al*., 2012) and second-end DNA capture (Nimonkar *et al*., 2009). Rad52 has also been shown to contain two DNA-binding domains: the first binds ssDNA and the second binds either ssDNA or dsDNA (Arai *et al*., 2011). In contrast, DdrA has been conclusively demonstrated to be incapable of binding dsDNA (Harris *et al*., 2004). This fundamental difference between DdrA and Rad52 suggests that while DdrA should be assessed for its ability to promote Rad52-like ssDNA transactions, the possibility of a considerable degree of functional dissimilarity should not be discounted.

Further functional studies could also be aimed at examining the role of the C-terminal tail of DdrA. Our data suggest that the presence of the C-terminal tail slows down the kinetics of annealing *in vitro* and yet a gene encoding $DdrA^{1-157}$ has been shown to be incapable of conferring radioresistance to cells *in vivo.* Taken together, these findings strongly suggest that the C-terminal domain is important for reasons unrelated to ssDNA annealing. Our hypothesis posits that the tail serves as a scaffold for certain uncharacterized *in vivo* interactions. Pull-down assays, performed in conjunction with mass spectrometry, could be conducted in an effort to test this hypothesis. Tagged DdrA could be mixed with lysate of *D. radiodurans* cells exposed to a DNA damaging stimulus, such as MMC. Following immobilization, matrix-assisted laser desorption/ionization (MALDI) mass spectroscopy could be used to identify the potential binding partners of DdrA. The experiment could be repeated using truncated DdrA and/or mutant DdrA incapable of annealing.

4.5 Conclusion

Prior to work completed here, DdrA was only known to bind ssDNA. In this thesis, we have shown that DdrA is capable of ssDNA annealing, allowing for the classification of DdrA into the Rad52 superfamily of single-strand annealing proteins. Our work identified residues (K22 and K105) required for annealing and demonstrated that ssDNA annealing by DdrA is important for *Deinococcal* DNA damage repair *in vivo*. While our understanding of DdrA has now expanded, much is left unanswered. In particular, how does DdrA (and related Rad52 homologues) accurately anneal ssDNA, and why does *Deinococcus* require three distinct repair pathways in order to maintain its extraordinary radioresistance to DNA damage?

References

Anderson, A.W., Nordan, H.C., Cain, R.F., Parish, G., & Duggan, D. (1956). Studies on a radioresistant *micrococcus*. I, Isolation, morphology, cultural characteristics and resistance to gamma radiation. *The Journal of Food Science and Technology*, *10*, 575-578.

Anderson, D.G., & Kowalczykowski, S.C. (1997). The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA in a χ-regulated manner. *Cell*, *90*(1), 77-86.

Arai, N., Kagawa, W., Saito, K., Shingu, Y., Mikawa, T., Kurumizaka, H., & Shibata, T. (2011). Vital roles of the second DNA-binding site of Rad52 protein in yeast homologous recombination. *Journal of Biological Chemistry*, *286*(20), 17607-17.

Arrange, A.A., Phelps, T.J., Benoit, R.E., Palumbo, A.V., & White, D.C. (1993). Bacterial sensitivity to UV light as a model for ionizing radiation resistance. *Journal of Microbiological Methods*, *18*(2), 127-136.

Battista, J.R. (1997). Against all odds: the survival strategies of *Deinococcus radiodurans*. *Annual Review of Microbiology*, *51*, 203-224.

Bentchikou, E., Servant, P., Coste, G., & Sommer, S. (2010). A major role of the RecFOR pathway in DNA double-strand break repair through ESDSA in *Deinococcus radiodurans*. *Public Library of Science: Genetics*, *6*(1), e1000774.

Bi, B., Rybalchenko, N., Golub, E.I., & Radding, C.M. (2004). Human and yeast Rad52 proteins promote DNA strand exchange. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(26), 9568-9572.

Bjelland, S., & Seeberg, E. (2003). Mutagenicity, toxicity and repair of DNA base damage induced by oxidation. *Mutation Research*, *531*, 37-80.

Blaškovičová, J., Sochr, J., Koutsogiannis, A., Diamantidou, D., Kopel, P., Adam, V., & Labuda, J. (2017). Detection of ROS generated by UV-C irradiation of CdS quantum dots and their effect on damage to chromosomal and plasmid DNA. *Electroanalysis*, *30*(4), 698-704.

Boling, M., & Setlow, J.K. (1966). The resistance of *Micrococcus radiodurans* to ultraviolet radiation: III. A repair mechanism. *Biochimica et Biophysica Acta*, *123*(1), 26-33.

Bonura, T., Smith, K.C., & Kaplan, H.S. (1975). Enzymatic induction of DNA double-strand breaks in gamma-irradiated *Escherichia coli* K-12. *Proceedings of the National Academy of Sciences of the United States of America*, *72*(11), 4265-4269.

Bradner, W. (2001). Mitomycin C: a clinical update. *Cancer Treatment Reviews*, *27*(1), 35-50.

Brooks, B.W., & Murray, R.G.E. (1981). Nomenclature for "*Micrococcus radiodurans*" and other radiation-resistant *cocci*: *Deinococcaceae* fam. nov. and *Deinococcus* gen. nov., including five species. *International Journal of Systematic Bacteriology*, *31*, 353-360.

Burrell, A.D., Feldschreiber, P., & Dean, C.J. (1971). DNA-membrane association and the repair of double breaks in X-irradiated *Micrococcus radiodurans*. *Biochimica et Biophysica Acta*, *247*(1), 38-53.

Chang, X., Yang, L., Zhao, Q., Fu, W., Chen, H., Qiu, Z., Chen, J.A., Hu, R., & Shu, W. (2010). Involvement of RecF in 254 nm ultraviolet radiation resistance in *D. radiodurans* and *E. coli*. *Current Microbiology*, *61*(5), 458-64.

Crooke, S.T., & Bradner, W.T. (1976). Mitomycin C: a review. *Cancer Treatment Reviews*, *3*(3), 121-39.

Dalle-Donne, I., Rossi, R., Colombo, R., Giustarini, D., & Milzani, A. (2006). Biomarkers of oxidative damage in human disease. *Clinical Chemistry*, *52*(4), 601-23.

Daly, M.J. (2009). A new perspective on radiation resistance based on *Deinococcus radiodurans*. *Natural Reviews: Microbiology*, *7*(3), 237-45.

Daly, M.J., Gaidamakova, E.K., Matrosova, V.Y., Kiang, J.G., Fukumoto, R., Lee, D., Wehr, N.B., Viteri, G.A., Berlett, B.S., & Levine, R.L. (2010). Small-molecule antioxidant proteome-shields in *Deinococcus radiodurans*. *Public Library of Science One*, *5*(9), e12570.

Daly, M.J., Gaidamakova, E.K., Matrosova, V.Y., Vasilenko, A., Zhai, M., Leapman, R.D., Lai, B., Ravel, B., Li, S.W., Kemner, K.M., & Fredrickson, J.K. (2007). Protein oxidation implicated as the primary determinant of bacterial radioresistance. *Public Library of Science: Biology*, *5*(4), e92.

Daly, M.J., Gaidamakova, E.K., Matrosova, V.Y., Vasilenko, A., Zhai, M., Venkateswaran, A., Hess, M., Omelchenko, M.V., Kostandarithes, H.M., Makarova, K.S., Wackett, L.P., Fredrickson, J.K., & Ghosal, D. (2004). Accumulation of Mn(II) in *Deinococcus radiodurans* facilitates gamma-radiation resistance. *Science*, *306*(5698), 1025-8.

Daly, M.J., & Minton, K.W. (1996). An alternative pathway of recombination of chromosomal fragments precedes RecA-dependent recombination in the radioresistant bacterium *Deinococcus radiodurans*. *Journal of Bacteriology*, *178*(15), 4461-4471.

Daly, M.J., Ouyang, L., Fuchs, P., & Minton, K.W. (1994). *In vivo* damage and RecA-dependent repair of plasmid and chromosomal DNA in the radiation-resistant bacterium *Deinococcus radiodurans*. *Journal of Bacteriology*, *176*(12), 3508-3517.

De Bont, R., & van Larebeke, N. (2004). Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis*, *19*(3), 169-185.

Dillingham, M.S., & Kowalczykowski, S.C. (2008). RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiology and Molecular Biology Reviews*, *72*(4), 642-671.

Flint, D.H., Tuminello, J.F., & Emptage, M.H. (1993). The inactivation of Fe-S cluster containing hydro-lyases by superoxide. *Journal of Biological Chemistry*, *268*(30), 22369-76.

Grimsley, J., Masters, C., Clark, E., & Minton, K. (1991). Analysis by pulsed-field gel electrophoresis of DNA double-strand breakage and repair in *Deinococcus radiodurans* and a radiosensitive mutant. *International Journal of Radiation Biology*, *60*(4), 613-626.

Gutman, P.D., Carroll, J.D., Masters, C.I., & Minton, K.W. (1994). Sequencing, targeted mutagenesis and expression of a RecA gene required for the extreme radioresistance of *Deinococcus radiodurans*. *Gene*, *141*(1), 31-7.

Gutsche, I., Vujičić-Žagar, A., Siebert, X., Servant, P., Vannier, F., Castaing, B., Gallet, B., Heullin, T., de Groot, A., Sommer, S., & Sierre, L. (2008). Complex oligomeric structure of a truncated form of DdrA: a protein required for the extreme radiotolerance of *Deinococcus*. *Biochimica et Biophysica Acta*, *1784*(7-8), 1050-8.

Hansen, M.T. (1978). Multiplicity of genome equivalents in the radiation-resistant bacterium *Micrococcus radiodurans*. *Journal of Bacteriology*, *134*(1), 71-75.

Harris, D.R., Ngo, K.V., & Cox, M.M. (2008). The stable, functional core of DdrA from *Deinococcus radiodurans* R1 does not restore radioresistance *in vivo*. *Journal of Bacteriology*, *190*(19), 6475-82.

Harris, D.R., Tanaka, M., Saveliev, S.V., Jolivet, E., Earl, A.M., Cox, M.M., & Battista, J.R. (2004). Preserving genome integrity: the DdrA protein of *Deinococcus radiodurans* R1. *Public Library of Science: Biology*, *2*(10), e304.

Harsojo, Kitayama, S., & Matsuyama, A. (1981). Genome multiplicity and radiation resistance in *Micrococcus radiodurans*. *Journal of Biochemistry*, *90*(3), 877-80.

Honda, M., Okuno, Y., Yoo, J., Ha, T., & Spies, M. (2011). Tyrosine phosphorylation enhances Rad52-mediated annealing by modulating its DNA binding. *The European Molecular Biology Organization (EMBO) Journal*, *30*(16), 3368-3382.

Imlay, J.A. (2008). Cellular defenses against superoxide and hydrogen peroxide. *Annual Review of Microbiology*, *77*, 755-776.

Imlay, J.A. (2003). Pathways of oxidative damage. *Annual Review of Microbiology*, *57*, 395-418.

Iyer, L.M., Koonin, E.V., & Aravind, L. (2002). Classification and evolutionary history of the single-strand annealing proteins: RecT, Redβ, ERF and Rad52. *BioMed Central (BMC) Genomics*, *3*, 8.

Joshi, B.S., Schmid, R., Altendorf, K., & Apte, S.K. (2004). Protein recycling is a major component of post-irradiation recovery in *Deinococcus radiodurans* strain R1. *Biochemical and Biophysical Research Communications*, *320*(4), 1112-1117.

Kagawa, W., Arai, N., Ichikawa, Y., Saito, K., Sugiyama, S., Saotome, M., Shibata, T., & Kurumizaka, H. (2014). Functional analyses of the C-terminal half of the *Saccharomyces cerevisiae* Rad52 protein. *Nucleic Acids Research*, *42*(2), 941-51.

Kantake, N., Madiraju, M.V., Sugiyama, T., & Kowalczykowski, S.C. (2002). *Escherichia coli* RecO protein anneals ssDNA complexed with its cognate ssDNA-binding protein: A common step in genetic recombination. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(24), 15327-32.

Khairnar, N.P., Kamble, V.A., & Misra, H.S. (2008). RecBC enzyme overproduction affects UV and gamma radiation survival of *Deinococcus radiodurans*. *DNA Repair*, *7*(1), 40-47.

Kitayama, S., Asaka, S., & Totsuka, K. (1983). DNA double-strand breakage and removal of cross-links in *Deinococcus radiodurans*. *Journal of Bacteriology*, *155*, 1200-1207.

Kota, S., Kumar, C.V., & Misra, H.S. (2010). Characterization of an ATP-regulated DNA-processing enzyme and thermotolerant phosphoesterase in the radioresistant bacterium *Deinococcus radiodurans*. *Biochemical Journal*, *431*(1), 149-57.

Kowalczykowski, S.C., Dixon, D.A., Eggleston, A.K., Lauder, S.D., & Rehrauer, W.M. (1994). Biochemistry of homologous recombination in *Escherichia coli*. *Microbiological Reviews*, *58*(3), 401-65.

Kozmin, S.G., Sedletska, Y., Reynaud-Angelin, A., Gasparutto, D., & Sage, E. (2009). The formation of double-strand breaks at multiply damaged sites is driven by the kinetics of excision/incision at base damage in eukaryotic cells. *Nucleic Acids Research*, *37*, 1767-1777.

Krisko, A., & Radman, M. (2010). Protein damage and death by radiation in *Escherichia coli* and *Deinococcus radiodurans*. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(32), 14373-7.

Levin-Zaidman, S., Englander, J., Shimoni, E., Sharma, A.K., Minton, K.W., & Minsky, A. (2003). Ring-like structure of the *Deinococcus radiodurans* genome: a key to radioresistance? *Science*, *299*(5604), 254-6.

Liu, Y., Zhou, J., Omelchenko, M.V., Beliaev, A.S., Venkateswaran, A., Stair, J., Wu, L., Thompson, D.K., Xu, D., Rogozin, I.B., Gaidamakova, E.K., Zhai, M., Makarova, K.S., Koonin, E.V., & Daly, M.J. (2003). Transcriptome dynamics of *Deinococcus radiodurans* recovering from ionizing radiation. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(7), 4191-6.

Lloyd, R., & Buckman, C. (1985). Identification and genetic analysis of sbcC mutations in commonly used RecBC sbcB strains of *Escherichia coli* K-12. *Journal of Bacteriology*, *164*(2), 836-44.

Lok, B.H., & Powell, S.N. (2012). Molecular pathways: understanding the role of Rad52 in homologous recombination for therapeutic advancement. *Clinical Cancer Research*, *18*(23), 6400-6.

Mackay, V., & Linn, S. (1976). Selective inhibition of the DNase activity of the RecBC enzyme by the DNA binding protein from *Escherichia coli*. *The Journal of Biological Chemistry*, *251*(12), 3716-9.

Makarova, K.S., Aravind, L., Wolf, Y.I., Tatusov, R.L., Minton, K.W., Koonin, E.V., & Daly, M.J. (2001). Genome of the extremely radiation-resistant bacterium *Deinococcus radiodurans* viewed from the perspective of comparative genomics. *Microbiology and Molecular Biology Reviews*, *65*(1), 44-79.

Makharashvili, N., Koroleva, O., Bera, S., Grandgenett, D.P., & Korolev, S. (2004). A novel structure of DNA repair protein RecO from *Deinococcus radiodurans*. *Structure*, *12*(10), 1881-9.

Markillie, L.M., Varnum, S.M., Hradecky, P., & Wong, K.K. (1999). Targeted mutagenesis by duplication insertion in the radioresistant bacterium *Deinococcus radiodurans*: radiation sensitivities of catalase (katA) and superoxide dismutase (sodA) mutants. *The Journal of Bacteriology*, *181*(2), 666-9.

Mattimore, V., & Battista, J.R. (1996). Radioresistance of *Deinococcus radiodurance*: functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation. *The Journal of Bacteriology*, *178*(3), 633-637.

Mihandoost, E., Shirazi, A., Mahdavi, S.R., & Aliasgharzadeh, A. (2014). Can melatonin help us in radiation oncology treatments? *BioMed Research International*, *2014*, 578137.

Minton, K.W. (1994). DNA repair in the extremely radioresistant bacterium *Deinococcus radiodurans*. *Molecular Microbiology*, *13*(1), 9-15.

Minton, K.W., & Daly, M.J. (1995). A model for repair of radiation-induced DNA double-strand breaks in the extreme radiophile *Deinococcus radiodurans*. *Bioessays*, *17*(5), 457-64.

Misra, H.S., Khairnar, N.P., Kota, S., Shrivastava, S., Joshi, V.P., & Apte, S.K. (2006). An exonuclease I-sensitive DNA repair pathway in *Deinococcus radiodurans*: a major determinant of radiation resistance. *Molecular Microbiology*, *59*(4), 1308-16.

Moeller, R., Douki, T., Rettberg, P., Reitz, G., Cadet, J., Nicholson, W.L., & Horneck, G. (2010). Genomic bipyrimidine nucleotide frequency and microbial reactions to germicidal UV radiation. *Archives of Microbiology*, *192*, 521-529.

Moseley, B.E., & Evans, D.M. (1983). Isolation and properties of strains of *Micrococcus* (*Deinococcus*) *radiodurans* unable to excise ultraviolet light-induced

pyrimidine dimers from DNA: evidence for two excision pathways. *Microbiology*, *129*(8), 2437-2445.

Muskavitch, K.M., & Linn, S. (1982). A unified mechanism for the nuclease and unwinding activities of the RecBC enzyme of *Escherichia coli*. *The Journal of Biological Chemistry*, *257*, 2641-264.

Nimonkar, A.V., Sica, R.A., & Kowalczykowski, S.C. (2009). Rad52 promotes second-end DNA capture in double-stranded break repair to form complement-stabilized joint molecules. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(9), 3077-3082.

Norais, C.A., Chitteni-Pattu, S., Wood, E.A., Inman, R.B., & Cox, M.M. (2009). DdrB protein, an alternative *Deinococcus radiodurans* SSB induced by ionizing radiation. *The Journal of Biological Chemistry*, *284*, 21402-21411.

Pardo, B., Gómez-González, B., & Aquilera, A. (2009). DNA repair in mammalian cells: DNA double-strand break repair: how to fix a broken relationship. *Cellular and Molecular Life Sciences (CMLS)*, *66*(6), 1039-56.

Parsons, C.A., Baumann, P., Van Dyck, E., & West, S.C. (2000). Precise binding of single-stranded DNA termini by human Rad52 protein. *The EMBO Journal*, *19*(15), 4175-4181.

Pfeifer, G.P. (1997). Formation and processing of UV photoproducts: effects of DNA sequence and chromatin environment. *Photochemistry and Photobiology*, *65*, 270-283.

Potts, M. (1994). Desiccation tolerance of prokaryotes. *Microbiological Reviews*, *58*(4), 755-805.

Repar, J., Cvjetan, S., Slade, D., Radman, M., Zahradka, D., & Zahradka, K.

(2010). RecA protein assures fidelity of DNA repair and genome stability in *Deinococcus radiodurans*. *DNA Repair*, *9*(11), 1151-1161.

Roche, B., Aussel, L., Ezraty, B., Mandin, P., Py, B., & Barras, F. (2003). Iron/sulfur proteins biogenesis in prokaryotes: formation, regulation and diversity. *Biochimica et Biophysica Acta*, *1827*(3), 455-69.

Saotome, M., Saito, K., Yasuda, T., Ohtomo, H., Sugiyama, S., Nishimura, Y., Kurumizaka, H., & Kagawa, W. (2018). Structural basis of homology-directed DNA repair mediated by Rad52. *iScience*, *3*, 50-62.

Shashidhar, R., Kumar, S.A., Misra, H.S., & Bandekar, J.R. (2010). Evaluation of the role of enzymatic and nonenzymatic antioxidant systems in the radiation resistance of *Deinococcus*. *Canadian Journal of Microbiology*, *56*(3), 195-201.

Shukla, M., Chaturvedi, R., Tamhane, D., Vyas, P., Archana, G., Apte, S., Bandekar, J., & Desai, A. (2007). Multiple stress-tolerance of ionizing radiation-resistant bacterial isolates obtained from various habitats: correlation between stresses. *Current Microbiology*, *54*(2), 142-8.

Singleton, M.R., Wentzell, L.M., Liu, Y., West, S.C., & Wigley, D.B. (2002). Structure of the single-strand annealing domain of human Rad52 protein. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(21), 13492-7.

Slade, D., Lindner, A.B., Paul, G., & Radman, M. (2009). Recombination and replication in DNA repair of heavily irradiated *Deinococcus radiodurans*. *Cell*, *136*(6), 1044-1055.

Sweet, D.M., & Moseley, B.E.B. (1976). The resistance of *Micrococcus radiodurans* to killing and mutation by agents which damage DNA. *Mutation Research*, *34*, 175-186.

Sugiman-Marangos, S.N., Weiss, Y.M., & Junop, M.S. (2016). Mechanism for accurate, protein-assisted DNA annealing by *Deinococcus radiodurans* DdrB. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(16), 4308-13.

Tanaka, M., Earl, A.M., Howell, H.A., Park, M.J., Eisen, J.A., Peterson, S.N., & Battista, J.R. (2004). Analysis of *Deinococcus radiodurans*'s transcriptional response to ionizing radiation and desiccation reveals novel proteins that contribute to extreme radioresistance. *Genetics*, *168*(1), 21-33.

Tanaka, M., Narumi, I., Funayama, T., Kikuchi, M., Watanabe, H., Matsunaga, T., Nikaido, O., & Yamamoto, K. (2005). Characterization of pathways dependent on the uvsE, uvrA1, or uvrA2 gene product for UV resistance in *Deinococcus radiodurans*. *The Journal of Bacteriology*, *187*(11), 3693-7.

Tian, B., Wu, Y., Sheng, D., Zheng, Z., Gao, G., & Hua, Y. (2004). Chemiluminescence assay for reactive oxygen species scavenging activities and inhibition on oxidative damage of DNA in *Deinococcus radiodurans*. *Luminescence*, *19*(2), 78-84.

Timmins, J., Leiros, I., & McSweeney, S. (2007). Crystal structure and mutational study of RecOR provide insight into its mode of DNA binding. *The EMBO Journal*, *26*, 3260-3271.

Tsang, S.S., Muniyappa, K., Azhderian, E., Gonda, D.K., Radding, C.M., Flory, J., & Chase, J.W. (1985). Intermediates in homologous pairing promoted by RecA protein. Isolation and characterization of active presynaptic complexes. *The Journal of Molecular Biology*, *185*(2), 295-309.

Varghese, A.J., & Daly, R.S., 3rd. (1970). Excision of cytosine-thymine adduct from the DNA of ultraviolet-irradiated *Micrococcus radiodurans*. *Photochemistry*

and Photobiology, *11*(6), 511-7.

Weng, M., Zheng, Y., Jasti, V.P., Champeil, E., Tomasz, M., Wang, Y., Basu, A.K., & Tang, M. (2010). Repair of mitomycin C mono- and interstrand cross-linked DNA adducts by uvrABC: a new model. *Nucleic Acids Research*, *38*(20), 6976-6984.

White, O., Eisen, J.A., Heidelberg, J.F., Hickey, E.K., Peterson, J.D., Dodson, R.J., Haft, D.H., Gwinn, M.L., Nelson, W.C., Richardson, D.L., Moffat, K.S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M., Vamathevan, J.J., Lam, P., McDonald, L., Utterback, T., Zalewski, C., Makarova, K.S., Aravind, L., Daly, M.J., Minton, K.W., Fleischmann, R.D., Ketchum, K.A., Nelson, K.E., Salzberg, S., Smith, H.O., Venter, J.C., & Fraser, C.M. (1999). Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science*, *286*, 1571-1577.

Xu, G., Wang, L., Chen, H., Lu, H., Ying, N., Tian, B., & Hua, Y. (2008). RecO is essential for DNA damage repair in *Deinococcus radiodurans*. *The Journal of Bacteriology*, *190*(7), 2624-8.

Xu, W., Shen, J., Dunn, C.A., Desai, S., & Bessmann, M.J. (2001). The Nudix hydrolases of *Deinococcus radiodurans*. *Molecular Microbiology*, *39*(2), 286-90.

Zahradka, K., Slade, D., Bailone, A., Sommer, S., Averbeck, D., Petranović, M., Lindner, A.B., & Radman, M. (2006). Reassembly of shattered chromosomes in *Deinococcus radiodurans*. *Nature*, *443*(7111), 569-573.

Zhou, Q., Zhiang, X., Xu, H., Xu, B., & Hua, Y. (2007). A new role of *Deinococcus radiodurans* RecD in antioxidant pathway. *Federation of European Microbiological Societies (FEMS) Microbiology Letters*, *271*(1), 118-12.

Zimmerman, J.M., & Battista, J.R. (2005). A ring-like nucleoid is not necessary for radioresistance in the *Deinococcaceae*. *BMC Microbiology*, *5*, 17.

Appendix

List of Constructs Used

Proteins

Protein	Sequence (N-terminus to C-terminus)	Parameters*		Applications
		MW (Da)	pI	
DdrA$^{1-151}$ (*D. rad*)	**Polyhistidine-tag:** *HHHHHH* **Linker region:** *RSDITSLYKKAGL* **TEV Cleavage Site:** *ENLYFQG* **Body of Protein:** MKLSDVQKRLQAPFPAHTVSWK PAAFNAERTRALLLAHVDARAV QDRLDAVCPDDWSFEMEVVSG AEVPTVKGRLTVLGVTREDIGEA PEGSMAAYKAAASDAMKRCAV QFGIGRYLYDLPKQWADWDDA RRGPKHLPELPEWARPDHERT	**Cut:** 16,896.18 **Uncut:** 20,004.61	**Cut:** 6.08 **Uncut:** 6.87	Structural studies
DdrA$^{1-151}$ (*D. swu*)	*HHHHHHRSDITSLYKKAGLENL YFQG*MTYAEVKARLAAPFPEQ RVRWRAQQVSKDRRTAMMV AYIDSRTVMERLDDVCPDGW AFDVELLPGATLVMKGRLTVL GQTRCDVGLAGEGGEAATHK AATSDALKRCAVHFGIGRYLYD LPAHWAAWDDRLRAPVQPPT LPQWALPGSERT	**Cut:** 16,793.27 **Uncut:** 19,901.70	**Cut:** 8.52 **Uncut:** 8.83	Insoluble protein; no significant applications
DdrA$^{1-157}$ (*D. rad*)	*HHHHHHRSDITSLYKKAGLENL YFQG*MKLSDVQKRLQAPFPAH TVSWKPAAFNAERTRALLLAH VDARAVQDRLDAVCPDDWSF EMEVVSGAEVPTVKGRLTVLG VTREDIGEAPEGSMAAYKAAA SDAMKRCAVQFGIGRYLYDLP KQWADWDDARRGPKHLPELP EWARPDHERTPGGAHL	**Cut:** 17,428.78 **Uncut:** 20,537.21	**Cut:** 6.20 **Uncut:** 6.90	Structural studies

DdrA$^{1\text{-}157}$ (*D. swu*)	*HHHHHHRSDITSLYKKAGLENL YFQG*MTYAEVKARLAAPFPEQ RVRWRAQQVSKDRRTAMMV AYIDSRTVMERLDDVCPDGWA FDVELLPGATLVMKGRLTVLG QTRCDVGLAGEGGEAATHKAA TSDALKRCAVHFGIGRYLYDLP AHWAAWDDRLRAPVQPPTL PQWALPGSERTAGAQHV	**Cut:** 17,356.88 **Uncut:** 20,465.31	**Cut:** 8.52 **Uncut:** 8.83	Insoluble protein; no significant applications
DdrA$^{1\text{-}160}$ (*D. swu*)	*HHHHHHRSDITSLYKKAGLENLYF QG*MTYAEVKARLAAPFPEQRVR WRAQQVSKDRRTAMMVAYIDS RTVMERLDDVCPDGWAFDVELL PGATLVMKGRLTVLGQTRCDVG LAGEGGEAATHKAATSDALKRCA VHFGIGRYLYDLPAHWAAWDDR LRAPVQPPTLPQWALPGSERTA GAQHVLQM	**Cut:** 17,729.36 **Uncut:** 20,837.79	**Cut:** 8.52 **Uncut:** 8.83	Same as above
DdrA$^{1\text{-}161}$ (*D. rad*)	*HHHHHHRSDITSLYKKAGLENLY FQG*MKLSDVQKRLQAPFPAHT VSWKPAAFNAERTRALLLAHVD ARAVQDRLDAVCPDDWSFEME VVSGAEVPTVKGRLTVLGVTRE DIGEAPEGSMAAYKAAASDAM KRCAVQFGIGRYLYDLPKQWAD WDDARRGPKHLPELPEWARPD HERTPGGAHLVQAM **Note:** A version of this construct with a C-terminal polyhistidine-tag was also designed, whereby the body of the protein, as outlined above, had *HHHHHH* attached to the C-terminus	**Cut:** 17,858.31 **Uncut:** 20,966.74 **C-term tagged:** 18,681.16	**Cut:** 6.20 **Uncut:** 6.90 **C-term tagged:** 6.59	**N-term tagged:** Structural studies **C-term tagged:** Structural & functional studies
DdrA$^{1\text{-}167}$ (*D. swu*)	*HHHHHHRSDITSLYKKAGLENLYFQ G*MTYAEVKARLAAPFPEQRVRWRA QQVSKDRRTAMMVAYIDSRTVMER LDDVCPDGWAFDVELLPGATLVMK GRLTVLGQTRCDVGLAGEGGEAAT HKAATSDALKRCAVHFGIGRYLYDLP AHWAAWDDRLRAPVQPPTLPQWA LPGSERTAGAQHVLQMLDSLRTE	**Cut:** 17,858.31 **Uncut:** 20,966.74	**Cut:** 6.20 **Uncut:** 6.90	Insoluble protein; no significant applications

DdrA$^{1\text{-}167}$ (*D. rad*)	*HHHHHHRSDITSLYKKAGLENLYFQ G*MKLSDVQKRLQAPFPAHTVSWKP AAFNAERTRALLLAHVDARAVQDRL DAVCPDDWSFEMEVVSGAEVPTVK GRLTVLGVTREDIGEAPEGSMAAYK AAASDAMKRCAVQFGIGRYLYDLPK QWADWDDARRGPKHLPELPEWAR PDHERTPGGAHLVQAMEQLRYE	**Cut:** 18,677.20 **Uncut:** 21,785.63	**Cut:** 5.96 **Uncut:** 6.63	Structural studies
DdrA$^{1\text{-}188}$ (*D. rad*)	*HHHHHHRSDITSLYKKAGLENLY FQG*MKLSDVQKRLQAPFPAHT VSWKPAAFNAERTRALLLAHVD ARAVQDRLDAVCPDDWSFEME VVSGAEVPTVKGRLTVLGVTRED IGEAPEGSMAAYKAAASDAMK RCAVQFGIGRYLYDLPKQWAD WDDARRGPKHLPELPEWARPD HERTPGGAHLVQAMEQLRYELP EDLDLQREVYKHLKAALGS	**Cut:** 21,054.92 **Uncut:** 24,163.45	**Cut:** 5.88 **Uncut:** 6.48	Same as above
DdrA$^{1\text{-}190}$ (*D. swu*)	*HHHHHHRSDITSLYKKAGLENLYFQ G*MTYAEVKARLAAPFPEQRVRWRA QQVSKDRRTAMMVAYIDSRTVMER LDDVCPDGWAFDVELLPGATLMKG RLTVLGQTRCDVGLAGEGGEAATHK AATSDALKRCAVHFGIGRYLYDLPAH WAAWDDRLRAPVQPPTLPQWALP GSERTAGAQHVLQMLDSLRTELPSD TDQLREVYRHLKLALSVVGP	**Cut:** 21,134.23 **Uncut:** 24,242.67	**Cut:** 7.77 **Uncut:** 8.48	Insoluble protein; no applications
Full-length DdrA (*D. rad*)	*HHHHHHRSDITSLYKKAGLENLYFQ G*MKLSDVQKRLQAPFPAHTVSWK PAAFNAERTRALLLAHVDARAVQD RLDAVCPDDWSFEMEVVSGAEVP TVKGRLTVLGVTREDIGEAPEGSM AAYKAAASDAMKRCAVQFGIGRYL YDLPKQWADWDDARRGPKHLPEL PEWARPDHERTPGGAHLVQAME QLRYELPEDLDLQREVYKHLKAALG SIHPVPTGPVPTNPVQGGRAA	**Cut:** 23,002.15 **Uncut:** 26,110.58	**Cut:** 6.17 **Uncut:** 6.71	Structural & functional studies

Full-length DdrA (*D. swu*)	*HHHHHHRSDITSLYKKAGLENLYFQ* *G*MTYAEVKARLAAPFPEQRVRWRA QQVSKDRRTAMMVAYIDSRTVMER LDDVCPDGWAFDVELLPGATLVMK GRLTVLGQTRCDVGLAGEGGEAAT HKAATSDALKRCAVHFGIGRYLYDLP AHWAAWDDRLRAPVQPPTLPQWA LPGSERTAGAQHVLQMLDSLRTELP SDTDQLREVYRHLKLALSVVGPPED QDRALVAQ	**Cut:** 22,456.67 **Uncut:** 25,565.11	**Cut:** 6.52 **Uncut:** 7.33	Functional Studies
Full-length DdrA K22A/K105A (*D. rad*)	*HHHHHHRSDITSLYKKAGLENLYFQ* *G*MKLSDVQKRLQAPFPAHTVSW**A** PAAFNAERTRALLLAHVDARAVQD RLDAVCPDDWSFEMEVVSGAEVP TVKGRLTVLGVTREDIGEAPEGSM AAYKAAASDAMARCAVQFGIGRYL YDLPKQWADWDDARRGPKHLPEL PEWARPDHERTPGGAHLVQAME QLRYELPEDLDLQREVYKHLKAALG SIHPVPTGPVPTNPVQGGRAA	**Cut:** 22,887.96 **Uncut:** 25,996.39	**Cut:** 5.80 **Uncut:** 6.36	Functional Studies
Full-length Rad52 *(H. sapiens)*	MSGTEEAILGGRDSHPAAGGGSV LCFGQCQYTAEEYQAIQKALRQRL GPEYISSRMAGGGQKVCYIEGHRV INLANEMFGYNGWAHSITQQNV DFVDLNNGKFYVGVCAFVRVQLK DGSYHEDVGYGVSEGLKSKALSLE KARKEAVTDGLKRALRSFGNALGN CILDKDYLRSLNKLPRQLPLEVDLTK AKRQDLEPSVEEARYNSCRPNMA LGHPQLQQVTSPSRPSHAVIPADQ DCSSRSLSSSAVESEATHQRKLRQK QLQQQFRERMEKQQVRVSTPSAE KSEAAPPAPPVTHSTPVTVSEPLLE KDFLAGVTQELIKTLEDNSEKWAV TPDAGDGVVKPSSRADPAQTSDT LALNNQMVTQNRTPHSVCHQKP QAKSGSWDLQTYSADQRTTGNW ESHRKSQDMKKRKYDPS	**Cut:** 46,168.66	**Cut:** 8.49	Functional Studies

* Both parameters, the molecular weight and the isoelectric point (pI), were computed using ExPASy ProtParam.

DNA

Plasmids

Plasmid	Tag Encoded	Selection	Applications
pUC57	No fusions	Kanamycin	Entry vector for all constructs of DdrA
pDEST527	TEV His_6 tag	Ampicillin	Expression vector for all N-terminally polyhistidine-tagged constructs of DdrA
pDEST14	No fusions	Ampicillin	Expression vector for all C-terminally polyhistidine-tagged constructs of DdrA
pSF2285	SUMO His_6 tag	Ampicillin	Expression vector for full-length hRad52 (Obtained from Dr. Mauro Modesti)
pRAD1	No fusions	Chloramphenicol	Expression of proteins for survival assay

Oligonucleotides

Oligo	Sequence (5’ to 3’)	Applications
14 nt poly A	AAAAAAAAAAAAAA	**Unlabelled**: Co-crystallization **Labelled:** ssDNA binding assessment
14 nt poly T	TTTTTTTTTTTTTT	Same as above
21 nt poly A	AAAAAAAAAAAAAAAAAAAA	Same as above
21 nt poly T	TTTTTTTTTTTTTTTTTTTT	Same as above
28 nt poly A	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	Same as above
28 nt poly T	TTTTTTTTTTTTTTTTTTTTTTTTTTTT	Same as above
29 nt poly A	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	Co-crystallization
29 nt poly T	TTTTTTTTTTTTTTTTTTTTTTTTTTTTT	Same as above

30 nt poly A	AAAAAAAAAAAAAAAAAAAAAAAAAA	Same as above
30 nt poly T	TTTTTTTTTTTTTTTTTTTTTTTTTTTT	Same as above
35 nt poly A	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	**Unlabelled**: Co-crystallization **Labelled:** ssDNA binding assessment
35 nt poly T	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	Same as above
42 nt poly A	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	Same as above
42 nt poly T	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	Same as above
49 nt poly A	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	Same as above
49 nt poly T	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	Same as above
Mismatched DNA: Forward Strand	TGCT**T**GCT**T**GCT**T**GCT**T**GCT**T**GCT**T**GCT	Co-crystallization
Mismatched DNA: Forward Strand	AGC**T**AGC**T**AGC**T**AGC**T**AGC**T**AGC**T**AGCA	Same as above

Oligo 1	GCAATTAAGCTCTAAGCCATCCGCAAAAATGACCTCTTATCAAAAGGA	**Unlabelled**: Termination of ssDNA annealing **Labelled:** Commencement of ssDNA annealing
Oligo 2	TCCTTTTGATAAGAGGTCATTTTTGCGGATGGCTTAGAGCTTAATTGC	Commencement of ssDNA annealing

Curriculum Vitae

Curriculum Vitae

Education

Master of Science – *Biochemistry* **Sep, 2016 – Aug, 2019**

Western University, London, Ontario

- Routinely performed a wide range of biochemical techniques, including, but not limited to, protein purification, gel electrophoresis, crystallization…
- Supervised the progress of three undergraduate students in the lab
- Collected X-ray diffraction data at McMaster University
- Worked as a teaching assistant (see below) as well as an exam proctor

Bachelor of Science – *Pharmacology* **Sep, 2011 – Aug, 2015**

University of Toronto, Toronto, Ontario

- Studied the fundamentals of pharmacology, including pharmacokinetics and pharmacodynamics, as well as the fundamentals of biochemistry, such as genomics, proteomics and metabolomics
- Investigated acute myeloid leukemia (AML) as part of my fourth-year research project under the supervision of Dr. Leonardo Salmena
- Used my pharmacology acumen to help propagate sensible drug policy as a member of the Canadian Students for Sensible Drug Policy (CSSDP)

Work Experience

Teaching Assistant **Jan, 2018 – April, 2018**

Western University, London, Ontario

- Taught 3rd year undergraduate biochemistry students
- Helped elevate the students' writing skills through the revision of lab reports
- Taught the students fundamental biochemical techniques, such as bacterial transformation, cellular growth, spectrophotometry, DNA cloning, protein purification and statistical data analysis
- Helped students improve their scientific understanding and scientific writing by answering questions during office hours